

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2002, 10:12:21 ; Search time 7114.74 Seconds  
(without alignments)  
5903.176 Million cell updates/sec

Title: US-09-052-089A-7  
2007  
Sequence: 1 GTGCGGTGAGACGAATTGTG.....AAAAAAAAAAAAAAAAAAAA 2007

Scoring table: IDENTITY\_NUC  
Gapop 10.0', Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_pro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_ju:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pal:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
-----					

1	2007	100.0	2007	9	HSU77845	U77845 Human hTRIP
2	1891.2	94.2	2002	9	BC019283	BC019283 Homo sapi
3	1891.2	94.2	2010	9	BC000310	BC000310 Homo sapi
4	1890.8	94.2	2065	6	AR072729	AR072729 Sequence
5	1890.8	94.2	2065	6	BD006990	BD006990 Modulator
6	1087	54.2	2024	10	BC017374	BC017374 Mus muscu
7	1081	53.9	1981	10	MMU77844	U77844 Mus musculu
8	521.6	26.0	3694	10	BC006929	BC006929 Mus muscu
9	260	13.0	129747	9	HS465N24	AL031432 Human DNA
10	212.8	10.6	391	11	G23637	G23637 human SFS W
11	176.4	8.8	7542	6	AX350344	AX350344 Sequence
12	176.4	8.8	163424	2	AC099330	AC099330 Homo sapi
13	130	6.5	110000	2	AL353694_2	Continuation (3 of
14	130	6.5	163424	2	AC099330	AC099330 Homo sapi
15	74	3.7	39410	5	FR165X09	AJ010317 Fugu rubr
16	65.8	3.3	7218	6	166494	166494 Sequence 14
17	60.6	3.0	203984	2	AC067854	AC067854 Homo sapi
18	59.6	3.0	318930	2	AC073495	AC073495 Mus muscu
19	59	2.9	216191	2	AC090495	AC090495 Mus muscu
20	58.8	2.9	2065	14	AF410847	AF410847 Ovine her
21	58.8	2.9	239480	2	AC079818	AC079818 Mus muscu
22	58	2.9	133661	14	U93872	U93872 Kaposi's sa
23	57.8	2.9	110000	2	LMFLCHR34_16	Continuation (17 o
24	57.6	2.9	219200	10	AL589701	AL589701 Mouse DNA
25	57	2.8	3012	14	AF360120	AF360120 Human her
26	57	2.8	192675	10	AL589767	AL589767 Mouse DNA
27	56.4	2.8	224712	2	AL606969	AL606969 Mus muscu
28	56.2	2.8	174019	2	AC027267	AC027267 Homo sapi
29	56.2	2.8	185497	2	AC073446	AC073446 Homo sapi
30	56.2	2.8	223426	2	AL606962	AL606962 Mus muscu
31	56	2.8	282611	2	AL645746	AL645746 Mus muscu
32	55.8	2.8	28559	14	AF148805	AF148805 Kaposi's
33	55.4	2.8	3127	14	AF305694	AF305694 Kaposi's
34	55.4	2.8	207683	2	AC098712	AC098712 Mus muscu
35	55	2.7	35848	2	AC011553	AC011553 Homo sapi
36	54	2.7	110000	9	AC091454_2	Continuation (3 of
37	53.8	2.7	180385	9	AC007461	AC007461 Homo sapi
38	53.8	2.7	241432	10	AL589661	AL589661 Mouse DNA
39	53.4	2.7	195294	2	AC051623	AC051623 Mus muscu
40	53.4	2.7	202010	2	AC016961	AC016961 Homo sapi
41	53.4	2.7	236685	2	AC084744	AC084744 Mus muscu
42	53.2	2.7	894	9	ATRINVA	M25314 Owl monkey
43	53.2	2.7	1635	9	ATRINVA	M25313 Owl monkey
44	53.2	2.7	227605	2	AC073564	AC073564 Mus muscu
45	53	2.6	142277	2	AC016204	AC016204 Homo sapi

## ALIGNMENTS

RESULT 1

LOCUS	Human hTRIP (hTRIP) mRNA, complete cds.	2007 bp	linear	PRI 25-APR-1997
DEFINITION	U77845			
ACCESSION	U77845			
VERSION	U77845.1	GI:2039303		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.			
AUTHORS	Lee,S.Y., Lee,S.Y. and Choi,Y.			
TITLE	TRAF-Interacting protein (TRIP): a novel component of the tumor necrosis factor receptor (TNFR)- and CD30-TRAF signaling complexes that inhibits TRAF2-mediated NF-kappaB activation			
JOURNAL	J. Exp. Med. 185 (7), 1275-1285 (1997)			
REFERENCE MEDLINE	97258620			
AUTHORS	Lee,S.Y., Lee,S.Y. and Choi,Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (10-NOV-1996) Immunology, The Rockefeller University, 1230 York Avenue Box 295, New York, NY 10021, USA			

FEATURES  
Source Location/Qualifiers  
1..2007  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..2007  
/gene="htr1p"  
/gene="htr1p"  
103..1512  
/gene="htr1p"  
/codon\_start=1  
/product="htr1p"  
/protein\_id="AA05293.1"  
/db\_xref="GI:2039304"  
/translation="MPTRALCTICSDFFDHSRDVAIHCGHTFHLCILQSEETAEBS  
TCPOCRIOVQKRTIINKLFEDLAEEENVDREFLNEDNVNVAQLSQRKERSQV  
IIDTLRDLERNAIVVLSQALGKAEMLCSTLKQMKLEQOODETKOAEAGRLR  
SKMTKEOIELLOSLPEYEMIRDMGSGSAVEQLAYCVSLKKEYENIKERARAS  
GYADKLRKDESSRSKLIOTVSELDQALNELKSNKQDQSNADKEIMSLKRLTMQE  
TLMPPVASETDRLVLESPAPVEVNLKRRSFRODIDLNATFEDVDTTPARSSOH  
GYEKLCLERSHSPIQDVPKICKGPRKSQSLGQSQAGEDELVGAFPLFVRNA  
ILGOKPKRPRSSCSKDVVRFGFDLGRTKFIQPTVTMIRPLVRPKTKVKORV  
RYKTVPSLFOAKLDTFLMS"

BASE COUNT 517 a 518 c 558 g 414 t  
ORIGIN

Query Match 100.0%; Score 2007; DB 9; Length 2007;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCGGTGGAGCCGAATTTGAGCAGCGAGCGGGCCCTCTAGCAGCCGAGCTGT 60  
DB 1 GTGCGGTGGAGCCGAATTTGAGCAGCGAGCGGGCCCTCTAGCAGCCGAGCTGT 60  
QY 61 AGCAGTTCTTTGGCTGCTGGGCCCCCTTGATCCAGCCATCATCGCTATCGGCTGT 120  
DB 61 AGCAGTTCTTTGGCTGCTGGGCCCCCTTGATCCAGCCATCATCGCTATCGGCTGT 120  
QY 121 TGCATATCTGCTCCGACTTCTTGATCACTCCGCGAGCTGGCCGCTCATCGCTGCGC 180  
DB 121 TGCATATCTGCTCCGACTTCTTGATCACTCCGCGAGCTGGCCGCTCATCGCTGCGC 180  
QY 181 CACACCTTCCACTTGGCAGTGCCTTAATTCAGTCTTTGAGCAGCAGCAGCAAGTGGAGCTGC 240  
DB 181 CACACCTTCCACTTGGCAGTGCCTTAATTCAGTCTTTGAGCAGCAGCAGCAGCAAGTGGAGCTGC 240  
QY 241 CCACAGTGCAGATCCAGTGTGGGCAAAAGAACCATTTATCATTAAGCTCTTTGATCTT 300  
DB 241 CCACAGTGCAGATCCAGTGTGGGCAAAAGAACCATTTATCATTAAGCTCTTTGATCTT 300  
QY 301 GCCCAGAGGAGGAGAAATGTCTTGATCGAAATTTCTTAAGAAATGAATGACAAATGTC 360  
DB 301 GCCCAGAGGAGGAGAAATGTCTTGATCGAAATTTCTTAAGAAATGAATGACAAATGTC 360  
QY 361 AGAGCCAGCTTTCCAGAAAGACAGAGGAGAAACGACAGCCAGGTCATCATGACACT 420  
DB 361 AGAGCCAGCTTTCCAGAAAGACAGAGGAGAAACGACAGCCAGGTCATCATGACACT 420  
QY 421 CTGCGGGATCGCTGGAAGAACCAATGCTACTGCTGATCTCTGACAGCGGCTTGGGC 480  
DB 421 CTGCGGGATCGCTGGAAGAACCAATGCTACTGCTGATCTCTGACAGCGGCTTGGGC 480  
QY 481 AAGCGGAGATGCTGCTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGAG 540  
DB 481 AAGCGGAGATGCTGCTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGAG 540  
QY 541 GATGAGACCAACAGCAGCAGAGAGGCGGCGCTCGAGAGCAAGATGAAGACCATG 600  
DB 541 GATGAGACCAACAGCAGCAGAGAGGCGGCGCTCGAGAGCAAGATGAAGACCATG 600  
QY 601 GACCAATTTGAGCTTACTCCAGAGCCAGCTCCGAGGTGAGAGGATGATCCAGAGAC 660  
DB 601 GACCAATTTGAGCTTACTCCAGAGCCAGCTCCGAGGTGAGAGGATGATCCAGAGAC 660

QY 661 ATGGGTGTGGGACATGCAGCGGTGGAACAGCTGGCTGTACTGTGTCTCTCAAGAAA 720  
DB 661 ATGGGTGTGGGACATGCAGCGGTGGAACAGCTGGCTGTACTGTGTCTCTCAAGAAA 720  
QY 721 GACTAGCAGATATTTAAAGAGCGCAGGAGGCTTCAGGGAGGTGCTGACAACTGAGG 780  
DB 721 GACTAGCAGATATTTAAAGAGCGCAGGAGGCTTCAGGGAGGTGCTGACAACTGAGG 780  
QY 781 AAGGATTTGTTTTCTCCAGAACCAAGTTGACAGCTGCTCTGAATTTGGATCGAGGC 840  
DB 781 AAGGATTTGTTTTCTCCAGAACCAAGTTGACAGCTGCTCTGAATTTGGATCGAGGC 840  
QY 841 AAGTTAAGATGAAGTCAGCCAGAGGACTTACAGAGTGTGACAGGAATAATCATGAGC 900  
DB 841 AAGTTAAGATGAAGTCAGCCAGAGGACTTACAGAGTGTGACAGGAATAATCATGAGC 900  
QY 901 CTGAAAAAGAGCTAAACGATGCTGACAGGAACCTTGAACCTGACCAAGTGGCAATGAG 960  
DB 901 CTGAAAAAGAGCTAAACGATGCTGACAGGAACCTTGAACCTGACCAAGTGGCAATGAG 960  
QY 961 ACTGTGACCGGCTGGGTTTTAGAGAGCCAGCCCTGTGGAGGTGAATCGAAGCTCCGC 1020  
DB 961 ACTGTGACCGGCTGGGTTTTAGAGAGCCAGCCCTGTGGAGGTGAATCGAAGCTCCGC 1020  
QY 1021 CGGCATTCCTCCGATGATATTTGATCTCAATGCTACCTTTGATGTGATACCTCCCA 1080  
DB 1021 CGGCATTCCTCCGATGATATTTGATCTCAATGCTACCTTTGATGTGATACCTCCCA 1080  
QY 1081 GCGCGGCGCTCCAGCTGCCAGCATGTGTTACTACGAAAACTTTGCTTACAGAACTACAC 1140  
DB 1081 GCGCGGCGCTCCAGCTGCCAGCATGTGTTACTACGAAAACTTTGCTTACAGAACTACAC 1140  
QY 1141 TCCCAATTCAGGATGTCGCCAAGAAATATGCAAAAGGCCCAAGAGAGTCCAGCTC 1200  
DB 1141 TCCCAATTCAGGATGTCGCCAAGAAATATGCAAAAGGCCCAAGAGAGTCCAGCTC 1200  
QY 1201 TCACCTGGGTGGCCAGAGCTGTGCAGAGAGCCAGATGAGAACTGTTGTTGCTTCCCT 1260  
DB 1201 TCACCTGGGTGGCCAGAGCTGTGCAGAGAGCCAGATGAGAACTGTTGTTGCTTCCCT 1260  
QY 1261 ATTTTGTCCGGAATGCCATCTTACAGGCAAGAAACAGCCCAAGAGGCTCAGAGTGC 1320  
DB 1261 ATTTTGTCCGGAATGCCATCTTACAGGCAAGAAACAGCCCAAGAGGCTCAGAGTGC 1320  
QY 1321 TCTTGCAGAAAGATGTGGTAAAGACAGGCTTGATGGGCTGGTGGCCGCAAAATTC 1380  
DB 1321 TCTTGCAGAAAGATGTGGTAAAGACAGGCTTGATGGGCTGGTGGCCGCAAAATTC 1380  
QY 1381 ATCCAGCTACTGACACAGTCAATGATCCGCCCATTTGCTTAAAGCCCAAGACCAAGTT 1440  
DB 1381 ATCCAGCTACTGACACAGTCAATGATCCGCCCATTTGCTTAAAGCCCAAGACCAAGTT 1440  
QY 1441 AAGCAGAGGTGAGGTTGAAGACCGTGTCTTCTTCCAGGCCAAAGTGGACACTTC 1500  
DB 1441 AAGCAGAGGTGAGGTTGAAGACCGTGTCTTCTTCCAGGCCAAAGTGGACACTTC 1500  
QY 1501 CTGTGTGCTGAGGAACAGTGAAGTCTGACCAATGGCCAGACACATGCTCAACTGTAG 1560  
DB 1501 CTGTGTGCTGAGGAACAGTGAAGTCTGACCAATGGCCAGACACATGCTCAACTGTAG 1560  
QY 1561 TCAAGGACTGTCCAGCAGAGGTTTGTGACAGAGCCCTACTTTCGGGACACAGCTGAGGT 1620  
DB 1561 TCAAGGACTGTCCAGCAGAGGTTTGTGACAGAGCCCTACTTTCGGGACACAGCTGAGGT 1620  
QY 1621 GTAAGGGCAGACAAACAGGTGAGGTGAGTGTACACCCAGAGACTGCTTCTTCCTG 1680  
DB 1621 GTAAGGGCAGACAAACAGGTGAGGTGAGTGTACACCCAGAGACTGCTTCTTCCTG 1680  
QY 1681 CACCCGCGCCCACTCTCTAGAGCTGGAGGTGACATACCAAGCCCATGATCTCTGACGA 1740  
DB 1681 CACCCGCGCCCACTCTCTAGAGCTGGAGGTGACATACCAAGCCCATGATCTCTGACGA 1740  
QY 1741 GGTCTGCTCTGTGTCAGAGCTCTTGTATAGCCATGATCAGATGTGCTCAGACTCTTT 1800

Db	1741	GGTCTGTGTCGTGGCCAGGCTCTTGTTTATAGCCATGATCGATGTGGTCACAGCTCTTT	1800
Qy	1801	CTGGGCTTGGAGACACAGGTCACCTTTGTTGACTGTCTCTGTGGACACAGATGTCTTGAAGCA	1860
Db	1801	CTGGGCTTGGAGACACAGGTCACCTTTGTTGACTGTCTCTGTGGACACAGATGTCTTGAAGCA	1860
Qy	1861	TCTAGGAGGCTCAGGCCCAAGCTTCTACACGTCCTTTGACTTGGCTTGTACATATAGCCTGG	1920
Db	1861	TCTAGGAGGCTCAGGCCCAAGCTTCTACACGTCCTTTGACTTGGCTTGTACATATAGCCTGG	1920
Qy	1921	GCCAAAGCAGGGTGGGGAATGAGGATATAGACATGGGATCTATGGAGACAGATGGAAATTTT	1980
Db	1921	GCCAAAGCAGGGTGGGGAATGAGGATATAGACATGGGATGATGAGAGAGATGGAAATTTT	1980
Qy	1981	CCCCGAAAAAAAAAAAAAAAAAAAAA 2007	
Db	1981	CCCCGAAAAAAAAAAAAAAAAAAAAA 2007	

RESULT	2	
LOCUS	BC019283	
DEFINITION	BC019283 Homo sapiens, TRAF interacting protein, clone MGC:3999	
ACCESSION	BC019283	
VERSION	BC019283.1	GI:17939476
KEYWORDS	MGC.	
SOURCE	human.	

REFERENCE 1 (bases 1 to 2002)

JLLE  
JOURNAL

USA  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK	MIN-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk

Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: DCDP/DRP  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@bcsc.bc.ca

Steven Jonker, Jennifer Asano, Ian Bosdet, Yaron Buterfield, Susanna Chan, Readman Chin, Chris Fiehl, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krywinski, Reta Kutscher, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McEavey, Steven Ness, Pawan Pandoh, Anna-Maria Prabh, Parvaneh Saedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Najaf van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MOC clone distribution information can be found through the I.M.A.G.E. Consortium/INTL at: <http://image.llnl.gov>  
Series: IRAL Plate: 9 Row: 1 Column: 1  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q1: 5032194.

FEATURES  
Source

CDS

```

/codon_start=1
/product="TRAF interacting protein"
/protein_id="AAH19283.1"
/db_xref="GI:17939477"
/translation="MPTRALCTISDFPDHSRDVAAHICGTHFLQCLIQWFETAPSP
TCCPCRQIOWGKRTITINKLPDLAAEEENVLDAPLKLELNLRDLSQKKEKRRRL
IITDLRDLERNAATVYSLOQALCKAEKLELTKOMKTELOOODETRKOQDEARRSL
SKMTMEQITELLLOSREPEVEEMI RDMGVGSAAEQALAVTCSLKEEENLEKARKAS
GEVADKLPRKLDLSRSRSKIQTVSELDAKLELKAQNDLOADEIMSLKRLMTLOE
TLANLPVAASETVDRLVLESPAPVEVNKLKRPSPRDDIDLNATFDVTPPARPSSSO
GYKYLCLCKESHSPIDQVPKKICGKPRRESQSLSGGQSCAPEDELTGAPRIFVRNNA
ILGOKORPKRPRESSSCSKDQVVRITGFDLGSGTFRILQPIDTYMINRPLPVKPKRTKYQR
RKRVYPSLFOAKIDTFLMS"

```

Query Match	94.28;	Score 1891.2;	DB 9;	Length 2002;
Best Local Similarity	98.48;	Pred. No. 0;		
Matches 1964; Conservative	0;	Mismatches 23;	Indels 9;	Gaps 5;

QY	20 GAAGCAAGCGGAGGGGGCG --- CTCACGAAGCCGACCTAGCAGTTCTTTCG 75           
Db	1 GAAGCAAGCGGAGGGCGCGTGCTACGAAGCCGACCTAGCAGTTCTTTCG 60           

[illegible]

QY 136 GACTTCTTCGATCACTCCCCGAGCTGGCCGCCATCCACTGCGGCCACACCTTCCACTTG 195

Db 121 GACTTCTTCGATCACTCCCCGACGTGGCCGCCCATCCACTGGGGCCACACCTTCCACTTG 180

181 CACCTCCCTTAAATTCACCTCCCTTTCACCAACCAACCAACCTCCCAACCAACCTCCCAATC 240

256 CAGGTGGCAAAGAACCATTTATCAATAAGCTCTTCTTTGATCTTGCCAGGAGGAG 315

22 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 105

Db 301 AATGTCCTGGATGCAGAACTCTTAAGAATGAACCTGGACAACTGTCAGAGCCCAAGCTTCC 360

376 CAGAAAGACACGAGAAACGAGACAGCCAGGTCATCATCGACACTCTGGGGATACGCTG 435

[illegible]

Db 421 GAAGAACGCAATGCTACTGTGATCTCTGCAGCAGGCCCTGGGCAAGGCCGAGATGCTG 480

Q7 450 1G1CCACAC1GAAAAACAGAIUAA6IAC11ABAGCACGACGAGAIUAGALCAAACA 333  

556 GCACAAGAGGAGCGCGCGGCTCAGGAGCAAGATGAAGACCATGGAGCAGATTGAGCTT 615

Db 541 GCACAAGAGGAGGCCCGCGGCTCAGGAGCAAGATGAAGACCATTGGAGCAGATTGAGCTT 600

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

676 TCAGCGTGACACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAAGAGTACGAGAATCTA 735

DD 001 ICAGCGGCGAATCAGC1GGC1G1G1AC1G1G1G1C1C1CAAGAAAGAG1ACGAGAA1C1A / 20

Db 721 AAAGAGGCACGGAGGCTCAGGGGAGTGCTGACCAAGCTGAGGAAGATTGTTTCC 780

QY	796	TCACAACCAAGTTGCGAGACAGCTACTCTCGAATTTGGATATCAGGCCAAGTTAGACAGCAAG	855
Db	781	TCACAAACCAAGTTGCGAGACAGCTACTCTCGAATTTGGATATCAGGCCAAGTTAGACCTGAAG	840
QY	856	TCAGCCCAAGAGGACTTACACAGTGTCTGACACAGAGAAATCATGTAGCCCTGAAAAAAGACTA	915
Db	841	TCAGCCCAAGAGGAACTTACACAGTGTCTGACACAGAGAAATCATGTAGCCCTGAAAAAAGACTA	900
QY	916	ACGATGCTGCGAGGAACCTTGTAACCTGCCACAGTGGCCAGTGAAGACTGCGACGCCCTG	975
Db	901	ACGATGCTGCGAGGAACCTTGTAACCTGCCACAGTGGCCAGTGAAGACTGCGACGCCCTG	960
QY	976	GTTTATAGAGACCCAGGCCCTGTGGAGGTAAATCTGAAGCTCCGGCGCATCTCTTCCT	1033
Db	961	GTTTATAGAGACCCAGGCCCTGTGGAGGTAAATCTGAAGCTCCGGCGCATCTCTTCCT	1020
QY	1036	GATGATATTGATCTCAATGCTACTCTTGATGTGGATACTCCCCAGCCGCCCTCTCAGC	1099
Db	1021	GATGATATTGATCTCAATGCTACTCTTGATGTGGATACTCCCCAGCCGCCCTCTCAGC	1088
QY	1096	TCCCAGCATGTGTACTACAGAAAACTTTGGCTTAGAGAAGTACACCTCCCAATTGAGAT	1155
Db	1081	TCCCAGCATGTGTGTACTACAGAAAACTTTGGCTTAGAGAAGTACACCTCCCAATTGAGAT	1140
QY	1156	GTCCCCAAGAGATATGTCAAAAGGCCCCAGAGAGAGTCCAGCTCTACTGGGTGCCAG	1213
Db	1141	GTCCCCAAGAGATATGTCAAAAGGCCCCAGAGAGAGTCCAGCTCTACTGGGTGCCAG	1200
QY	1216	AGCTGTCCAGAGAGACCAAGATGAGAACTGTGTGGCTCTCCCTATTTTGTGCCGAAAT	1279
Db	1201	AGCTGTCCAGAGAGACCAAGATGAGAACTGTGTGGCTCTCCCTATTTTGTGCCGAAAT	1266
QY	1276	GCCATTCCTTAGCCCAAGAAACAGCCCCAAAGGCCCCAGGTCAAGAGTCTCTTGACAAAGAT	1335
Db	1261	GCCATTCCTTAGGCCCAAGAAACAGCCCCAAAGGCCCCAGGTCAAGAGTCTCTTGACAAAGAT	1322
QY	1336	GTGTTAAGAGACAGCTTCGATGGGCTGTGGTGGCCGGACAAATTCATCCAGCTACTGAC	1399
Db	1321	GTGTTAAGAGACAGCTTCGATGGGCTGTGGTGGCCGGACAAATTCATCCAGCTACTGAC	1386
QY	1396	ACAGTCATGATCCGCCCATATGGCCGTGTAAACCCCAAGACCAAGTTTAAGACAGAGGGTAGG	1455
Db	1381	ACAGTCATGATCCGCCCATATGGCCGTGTAAACCCCAAGACCAAGTTTAAGACAGAGGGTAGG	1440
QY	1456	GTGAAGACCGTGCCTTCTCTTCACAGGCCCAAGCTGGACACTTCCCTGTGTGTGTAGAGA	1515
Db	1441	GTGAAGACAGTGCCTTCTCTTCACAGGCCCAAGCTGGACACTTCCCTGTGTGTGTAGAGA	1500
QY	1516	CAGTGTAGCTTACCAATGTGGCCAGACACATGCTCTGCAACTTTTAGTCAAGACTGTCCAG	1573
Db	1501	CAGTGTAGCTTACCAATGTGGCCAGACACATGCTCTGCAACTTTTAGTCAAGACTGTCCAG	1560
QY	1576	GCAGGG--TTTGTGTGACAGAGCCCTACTTTCGGGACAGCTGAGGTGTAAAGGGAGACA	1633
Db	1561	GCAGGGGTTTGTGTGACAGAGCCCTACTTTCGGGACAGCTGAGGTGTAAAGGGAGACA	1620
QY	1634	AACAGGTGAGGTTAGTGTACACCCAGAGACTGCTCTTCTGACCCTCAACCTGGCCAC	1693
Db	1621	AACAGGTGAGGATGATGTTGTGGACACCCAGAGACTGCTCTTCTGACCCTCAACCTGGCCAC	1680
QY	1694	TCCTTACAGTGTGGAGCTGACATGACACGCCCACTGATCTCTGTACAGAGGTCTCTGCT	1752
Db	1681	TCCTTACAGTGTGGAGCTGACATGACACGCCCACTGATCTCTGTACAGAGGTCTCTGCT	1740
QY	1753	TTTGGCAGGCTCTGTTTATATAGCCATGATCAATGATGTGTGACACTCTTCTGTGGCCGTGAG	1812
Db	1741	TTTGGCAGGCTCTGTTTATATAGCCATGATCAATGATGTGTGACACTCTTCTGTGGCCGTGAG	1800
QY	1813	ACCAAGGTCATTTGTGACTGTCTGTGTGACACAGAGTCTTGAGAGACTCTCAGGCACGC	1877
Db	1801	ACCAAGGTCATTTGTGACTGTCTGTGTGACACAGAGTCTTGAGAGACTCTCAGGCACGC	1866
QY	1873	TCAGCCCAAGCTTCTACCTGCCTTTGACTTGTCTTA--GCATAGCTGTGGCCCAAGCAAGG	1931

Accession	Sequence	Length
Db 1861	TCAGCCCAAGCTTTACCTGCTGTCCTTCCTTCAGGCATBGCCTGGGCCAACAGAGG	19220
Qy 1932	TGGGGAAATGAGGATAGACATGGGATGTATGAGAGAGATGGAAGATTTTCCCAAAAAA	1991
Db 1921	TGGGTAATGAGGAGATAG-CATGGATGTATGAGAGAGATGGAAGATTTTCATGTAAATA	1979
Qy 1992	AAAAAAAAAAAAAAAAA 2007	
Db 1980	AAATTAAAAAAAAAAAA 1995	

RESULT	3
BC000310	BC000310
LOCUS	Homo sapiens, TRAF interacting protein, clone MGC:8424
DEFINITION	IMAGE:2821007, mRNA, complete cds.
ACCESSION	BC000310
VERSION	BC000310.1
KEYWORDS	MGC.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

I (bases 1 to 2010)  
Strausberg, R.  
Direct Submission  
Submitted (15-NOV-2000) National Institutes of Health, Mammalian

**JOURNAL** Submitted (15-NOV-2000) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCRD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULIN),  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland.  
Web site: <http://www.nisc.nih.gov/>

Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mcgen@nih.nih.gov](mailto:nisc.mcgen@nih.nih.gov)  
 Shenjiao X., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lin, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
 McEwen, J., Pearson, R., Snyder, B., Statulonis, S., Thomas, P.J.,  
 Tongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 1 Row: e Column: 7  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 50332194.

## FEATURES

### SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="locusid:10293"
/culture="MGC:8424 IMAGE:2821007"
/culture_type="Lung, small cell carcinoma"
/clone_id="NH_MGC_7"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
97..1506

```

```

/codon_start=1
/product="TRAF interacting protein"
/protein_id="AAH0310.1"
/db_xref="GI:12653039"
/translation="MPTALCTCTCDSDFDHSRDVAAIHCGHTFHLQCLIQWETAPSR
TIDPCRIQVGKRTIINKLPDLAOGEEENVDAEELKINELNVAQLDSQKQEKESDQV
IIDLRLTERLNATVYSLDQALGKAEMLSLTKQKKRYLEQOODETRKQOEARRR
```

```

BASE COUNT      517 a      519 c      561 g      413 t
ORIGIN

```

Query Match	94.2%;	Score 1891.2;	DB 9;	Length 2010;
Best Local Similarity	98.4%;	Pred. No. 0;		
Matches 1964; Conservative	0;	Mismatches 23;	Indels 9;	Gaps 5;

Qy	20	GAAGCAAGCGAAGCGGGGGCG---CTCTACGAAGCCGGACCTGTAGACAGTTCTTTGGC	75
Dp	10	GAAGCAAGCGAAGCGGGGGGGCGCTGTACGAAGCCGGACCTGTAGACAGTTCTTTGGC	69
Qy	76	TGCCGTGGGCCCCCTTGAGTCCAGGCATATGCTTTCCGTCCTGTGACCATATCTGGTCC	135
Dp	70	TGCCGTGGGCCCCCTTGAGTCCAGGCATATGCTTTCCGTCCTGTGACCATATCTGGTCC	129
Qy	136	GACATTCTTGATCTACTCCCGCGAGCTGGCGCCATCCATCGCGGCCACACCTTCCACTTG	195
Dp	130	GACTTCTTGATCTACTCCCGCGAGCTGGCGCCATCCATCGCGGCCACACCTTCCACTTG	189
Qy	196	CAGTGCCTAATTCAGTCCCTTTGACACAGCCAAAGTGGAGCTGGACCTGCCACAGTCCGAATC	255
Dp	190	CAGTGCCTAATTCAGTGGTTTGAACACAGCCAAAGTGGAGCTGGACCTGCCACAGTCCGAATC	249
Qy	256	CAGTTTGGCAAAAGAACCCATTATCAATTAAGCTCTTCTTTATCTTTGGCCAGAGAGAGAG	315
Dp	250	CAGTTTGGCAAAAGAACCCATTATCAATTAAGCTCTTCTTTATCTTTGGCCAGAGAGAGAG	309
Qy	316	AATGCTTGATTCGAGAAATTTCTTAAAGATGAACCTGGACAAATGTGACAGCCACAGCTTCC	375
Dp	310	AATGCTTGATTCGAGAAATTTCTTAAAGATGAACCTGGACAAATGTGACAGCCACAGCTTCC	369
Qy	376	CAGAAAGCAAGAGAGAAACGAGACAGCCAGGTCATCTACCAACTCTGCGGGATAGCGTG	435
Dp	370	CAGAAAGCAAGAGAGAAACGAGACAGCCAGGTCATCTACCAACTCTGCGGGATAGCGTG	429
Qy	436	GAAGAACGCATCTACTGTGTGTATCTCTGACAGCAGGCCCTTTGGCAAGCCCGAGATCGT	495
Dp	430	GAAGAACGCATCTACTGTGTGTATCTCTGACAGCAGGCCCTTTGGCAAGCCCGAGATCGT	489
Qy	496	TGCTTCACACTGAGAAAAGAGACATGACATCTTAGCAGCAGCAGAGTGTGAGACAAACAA	555
Dp	490	TGCTTCACACTGAGAAAAGAGACATGACATCTTAGCAGCAGCAGAGTGTGAGACAAACAA	549
Qy	556	GCACAAAGAGAGGCGGCGGCGCTCAGAGGCAAGATGGAAGCAATGTGAGCATTTGAGCTT	615
Dp	550	GCACAAAGAGAGGCGGCGGCGCTCAGAGGCAAGATGGAAGCAATGTGAGCATTTGAGCTT	609
Qy	616	CTACTCCAGACCCAGCCTCCCTGAGAGTGGAGAGATGATCCGAGACATGGGTGTGGGACAG	675
Dp	610	CTACTCCAGACCCAGCCTCCCTGAGAGTGGAGAGATGATCCGAGACATGGGTGTGGGACAG	669
Qy	676	TCAGCGGTGTGAACAGCTGGCGCTGTACTGTGTGTCTCTCAAGAAAGATACGAGAAATCTA	735
Dp	670	TCAGCGGTGTGAACAGCTGGCGCTGTACTGTGTGTCTCTCAAGAAAGATACGAGAAATCTA	729
Qy	736	AAAAGAGCAGCGAAGGCTCAGGGGGAGGTGGCTGACAAAGCTGAGGAAGATTTTGTTTCC	795
Dp	730	AAAAGAGCAGCGAAGGCTCAGGGGGAGGTGGCTGACAAAGCTGAGGAAGATTTTGTTTCC	789
Qy	796	TCCAGAAAGCAGTTGTCAGACAGTGTACTCTGGAATTTGGATAGAGCCAAAGTATAGAAGCAG	855
Dp	790	TCCAGAAAGCAGTTGTCAGACAGTGTACTCTGGAATTTGGATAGAGCCAAAGTATAGAAGCAG	849
Qy	856	TCAGCCCAAGAAAGACTTACAGAGTGTCTGACAAAGGAATTCATGAAGCCTGAAAAGAAAGCTA	915
Dp	850	TCAGCCCAAGAAAGACTTACAGAGTGTCTGACAAAGGAATTCATGAAGCCTGAAAAGAAAGCTA	909

QY	916	ACGATGCTGCAGGAAACCTTTGAACCTGACCACAGTGGCCAGTGTAGACTGTGACCGCTG	975
Db	910	ACGATGCTGCAGGAAACCTTTGAACCTGACCACAGTGGCCAGTGTAGACTGTGACCGCTG	969
QY	976	GTTTATGAGAGCCCCAGCCCGTGTGGAGGTGAATCTGAACCTCGCGGGCAGTCCTTCGCT	10353
Db	970	GTTTATGAGAGCCCCAGCCCGTGTGGAGGTGAATCTGAACCTCGCGGGCAGTCCTTCGCT	10289
QY	1036	GATGATATTGATCTCAATGCTACCTTTGATGTGATACTCGCCAGCCGGCCCTCCAGC	10955
Db	1030	GATGATATTGATCTCAATGCTACCTTTGATGTGATACTCGCCAGCCGGCCCTCCAGC	10899
QY	1096	TCCGAGCATGGTTCTTACGSAANAACCTTTGGCTGTAGAACTCACTACCTCCCAATTAGAT	11555
Db	1090	TCCGAGCATGGTTCTTACGSAANAACCTTTGGCTGTAGAACTCACTACCTCCCAATTAGAT	11499
QY	1156	GTCGCCAAGAAGATATCAAGAGGCCCCAGAGAAAGATGCCACCTGTCACCTGGTGGCCAG	12151
Db	1150	GTCGCCAAGAAGATATCAAGAGGCCCCAGAGAAAGATGCCACCTGTCACCTGGTGGCCAG	12099
QY	1216	AGCTGTGCAGGAGAGCCAGATGAGAACTGTGTGCTTCCCTATTTTGTCCGAAAT	12757
Db	1210	AGCTGTGCAGGAGAGCCAGATGAGAACTGTGTGCTTCCCTATTTTGTCCGAAAT	12699
QY	1276	GCCATTCCTAGGCCCGAAGAACAGCCCAAAAGGCCAGGTGACAGCTCTTCGACGCAAAAT	13353
Db	1270	GCCATTCCTAGGCCCGAAGAACAGCCCAAAAGGCCAGGTGACAGCTCTTCGACGCAAAAT	13299
QY	1336	GTCGTAAGACACGCTTCGATGGGGCTCGGTGGCCGGACAAAATTCAATCCAGCTTACTGAC	13955
Db	1330	GTCGTAAGACACGCTTCGATGGGGCTCGGTGGCCGGACAAAATTCAATCCAGCTTACTGAC	13899
QY	1396	ACAGCATGATCCGCCCATTTGCGCGTTAAGCCCAAGCAAGATTAAGAGAGGGTGGAG	14555
Db	1390	ACAGCATGATCCGCCCATTTGCGCGTTAAGCCCAAGCAAGATTAAGAGAGGGTGGAG	14499
QY	1456	GTGAAGACCGTCCCTTCTCTTCCAGGGCCAGCCTTGACACCTTCCCTGTGTGTCGTGAATA	15151
Db	1450	GTGAAGACCGTCCCTTCTCTTCCAGGGCCAGCCTTGACACCTTCCCTGTGTGTCGTGAATA	15099
QY	1516	CAGTGAGCTGTGACCAATGTGGCCAGACACATGTGCTGTGCAACTTTGTAGTCAAGACTGTCCAG	15757
Db	1510	CAGTGAGCTGTGACCAATGTGGCCAGACACATGTGCTGTGCAACTTTGTAGTCAAGACTGTCCAG	15699
QY	1576	GCAGGG - TTGTGTGAGACAGAGCCCTTACTTTGCGGAGCCAGCCTGAGAGGTGTAAAGGGACACA	16333
Db	1570	GCAGGGGTTTGTGTGAGACAGAGCCCCACTTTCGSGGACACCACTGTAGGTGTAAAGGGACACA	16299
QY	1634	AACAGGTGAGGGGTGAGTGTACACCCAGAGACTGTCTTCTGTGCCCCAC	16933
Db	1630	AACAGGTGAGGTGAGTGTGTGGACACCCAGAGACTGTCTTCTGTGCCCCAC	16899
QY	1694	TCCATACGATGGGAGCTGACATACCAAGCCCACTATGATCTCTGTACAGAGGTCTCTGCT - CTG	17522
Db	1690	TCCATACGATGGGAGCTGACATACCAAGCCCACTATGATCTCTGTACAGAGGTCTCTGCTCTG	17499
QY	1753	TTTGCCAGGCTCTTGTATTAAGCCATGATCAGATGTGTGAGACTCTTTCTGTGGCCCTGGAG	18122
Db	1750	TTTGCCAGGCTCTTGTATTAAGCCATGATCAGATGTGTGAGACTCTTTCTGTGGCCCTGGAG	18099
QY	1813	ACCAAGGTCACTTGTATACGTCTCTGTGTGACCAGAGTCTTGAGGCACTTCAGGACACC	18722
Db	1810	ACCAAGGTCACTTGTATACGTCTCTGTGTGACCAGAGTCTTGAGGCACTTCAGGACACC	18699
QY	1873	TCACGCCAAGCTTCTACCTCCCTTTTACCTGCTTCA - GCATAGCCTGGGCCCAACAGAG	19313
Db	1870	TCACGCCAAGCTTCTACCTCCCTTTTACCTGCTTCAAGCATAAGCCTGGGCCCAACAGAG	19299
QY	1932	TGGGGAATGAGGATAGACATGGGATGTATGAGAGGATGGAAGATTTTCCCGAAAAAA	19913
Db	1930	TGGGTAATGAGGATAG - CATGGGATGTATGAGAGGATGGAAGATTTTCACTGTAATA	19889

Qy 1992 AAAAAAAAAAAAAA 2007  
Db 1989 AAAAAAAAAAAAAA 2004

RESULT 4  
AR072729 2065 bp DNA linear PAT 28-AUG-2000  
LOCUS AR072729  
DEFINITION Sequence 1 from patent US 5948643.  
ACCESSION AR072729  
VERSION AR072729.1 GI:9999493  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2065)  
AUTHORS Rubinfeld,B., Polakis,P.G., Lingenfelter,C. and Vuong,T.T.  
TITLE Modulators of BRCA1 activity  
JOURNAL Patent: US 5948643-A 1 07-SEP-1999;  
FEATURES Location/Qualifiers  
source 1, 2065  
BASE COUNT 561 a 526 c 561 g 417 t  
ORIGIN

Query Match 94.2% Score 1890.8: DB 6: Length 2065;  
Best Local Similarity 98.9% Pred. No. 0:  
Matches 1946: Conservative 0: Mismatches 17: Indels 5: Gaps 4:

Qy 44 TACGAAGCCGAGACCTGATGCTTTGGCTGCTGGGCCCTTGAGTCAGCCATCA 103  
Db 44 TACGAAGCCGAGACCTGATGCTTTGGCTGCTGGGCCCTTGAGTCAGCCATCA 103

Qy 104 TGGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 163  
Db 104 TGGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 163

Qy 164 CCGCCTATCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 223  
Db 164 CCGCCTATCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 223

Qy 224 CACCAAGTGGAGCTGCGCCACAGTCCGCAATCCAGTGGCAAAAGAACCATATCAATA 283  
Db 224 CACCAAGTGGAGCTGCGCCACAGTCCGCAATCCAGTGGCAAAAGAACCATATCAATA 283

Qy 284 AGCTCTCTTTGATCTTGGCCAGAGAGAGAGAGATCTTTGGATCGAGATTTCTTAAAGA 343  
Db 284 AGCTCTCTTTGATCTTGGCCAGAGAGAGAGAGATCTTTGGATCGAGATTTCTTAAAGA 343

Qy 344 ATGAAGTGGACATGTCAGAGCCAGCTTCCAGAAAGCAAGAGAGAAACGAGACAGCC 403  
Db 344 ATGAAGTGGACATGTCAGAGCCAGCTTCCAGAAAGCAAGAGAGAAACGAGACAGCC 403

Qy 404 AGGTCAATCGACACTCTGCGGGATACGCTGGAGAACGCAATGCTACTGTGTATCTC 463  
Db 404 AGGTCAATCGACACTCTGCGGGATACGCTGGAGAACGCAATGCTACTGTGTATCTC 463

Qy 464 TGCAGAGAGCCCTTGGCAGAGCCGAGATGCTGTCTCCACTGGAAGAGAGAGAGTAACT 523  
Db 464 TGCAGAGAGCCCTTGGCAGAGCCGAGATGCTGTCTCCACTGGAAGAGAGAGAGTAACT 523

Qy 524 ACTTAGAGAGAGAGAGATGAGAACAAAGCAAGCAAGAGAGAGAGAGAGAGAGAGAG 583  
Db 524 ACTTAGAGAGAGAGAGATGAGAACAAAGCAAGCAAGAGAGAGAGAGAGAGAGAGAG 583

Qy 584 GCAAGATGAAGACATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 643  
Db 584 GCAAGATGAAGACATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 643

Qy 644 AGGAGATGATCCGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAG 703  
Db 644 AGGAGATGATCCGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGATGAG 703

Qy 704 GTGTGTCTCTCAAGAAAGATAGAGAAATCTAAAGAGCAGGAGGCTTCAGGGAGG 763  
Db 704 GTGTGTCTCTCAAGAAAGATAGAGAAATCTAAAGAGCAGGAGGCTTCAGGGAGG 763

Qy 764 TGGCTGACAAAGCTGAGAGAGATTTGTTTCTCCAGAAAGCAAGTTGACAGAGTCTACT 823  
Db 764 TGGCTGACAAAGCTGAGAGAGATTTGTTTCTCCAGAAAGCAAGTTGACAGAGTCTACT 823

Qy 824 CTGAATTTGATTCAGAGCCAAAGTTAGAACTGAACTGACGCCAGAGAGACTTACAGAGTCTG 883  
Db 824 CTGAATTTGATTCAGAGCCAAAGTTAGAACTGAACTGACGCCAGAGAGACTTACAGAGTCTG 883

Qy 884 ACAAGGAATCATGAGCTGCAAAAAGAGCTAATGCTGCGAGAGAAACCTTGAACCTCG 943  
Db 884 ACAAGGAATCATGAGCTGCAAAAAGAGCTAATGCTGCGAGAGAAACCTTGAACCTCG 943

Qy 944 CACCAAGTGGCAGTGAAGTGTGACCGGCTGGTTTGAAGAGCCAGGCCCTGTGGAGG 1003  
Db 944 CACCAAGTGGCAGTGAAGTGTGACCGGCTGGTTTGAAGAGCCAGGCCCTGTGGAGG 1003

Qy 1004 TGAATCTGAAGCTCGCGCGCCATCTTCCGTGATGATATTGATCTCAATGCTACCTTTG 1063  
Db 1004 TGAATCTGAAGCTCGCGCGCCATCTTCCGTGATGATATTGATCTCAATGCTACCTTTG 1063

Qy 1064 ATGTGATATCTCCCGCCAGCCCGGCTCCAGCTCCAGAGATGATCTACGAAAACTTT 1123  
Db 1064 ATGTGATATCTCCCGCCAGCCCGGCTCCAGCTCCAGAGATGATCTACGAAAACTTT 1123

Qy 1124 GCCTAGAGAAATGACACATCCCAATTCAGAGATGCTCCCAAGAAAGATATCAAAAGGCCCA 1183  
Db 1124 GCCTAGAGAAATGACACATCCCAATTCAGAGATGCTCCCAAGAAAGATATCAAAAGGCCCA 1183

Qy 1184 GGAAGAGTCCCAAGCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243  
Db 1184 GGAAGAGTCCCAAGCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243

Qy 1244 TGGTGTGCTCTCCCTATTTTGTGCGGAATGCTTCTAGGCCGGAAGAGCCCAAAA 1303  
Db 1244 TGGTGTGCTCTCCCTATTTTGTGCGGAATGCTTCTAGGCCGGAAGAGCCCAAAA 1303

Qy 1304 GGGCCAGTCAAGTCTCTTTCAGCAAGAGATGCTGTAAGAGACAGGCTTCGATGGGCTCG 1363  
Db 1304 GGGCCAGTCAAGTCTCTTTCAGCAAGAGATGCTGTAAGAGACAGGCTTCGATGGGCTCG 1363

Qy 1364 GTGGCCGGAACAAATTCATCCAGCTTACTGACACAGTCAATGATCCGCCATTTGCTGTTA 1423  
Db 1364 GTGGCCGGAACAAATTCATCCAGCTTACTGACACAGTCAATGATCCGCCATTTGCTGTTA 1423

Qy 1424 AGCCCAAGACCAAGGTTAAGCAAGAGGTTGAGGTTGAAGACCGTCTCTTCCAGG 1483  
Db 1424 AGCCCAAGACCAAGGTTAAGCAAGAGGTTGAGGTTGAAGACCGTCTCTTCCAGG 1483

Qy 1484 CCAAGCTGACACCTCTCTGTGTGTGAGAAACAGTGAAGTGTGACCAATGAGCCAGACACA 1543  
Db 1484 CCAAGCTGACACCTCTCTGTGTGTGAGAAACAGTGAAGTGTGACCAATGAGCCAGACACA 1543

Qy 1544 TGCCTGCAACTGTGATGTAAGAGTGTCCAGGAGGCTTTGTGTGAGAGAGCCCACT 1601  
Db 1544 TGCCTGCAACTGTGATGTAAGAGTGTCCAGGAGGCTTTGTGTGAGAGAGCCCACT 1601

Qy 1602 TTGCGGACAGGCTGAGGTGAAGGCAAGCAAGAGGAGAGGAGTGTGACAGCCAG 1661  
Db 1602 TTGCGGACAGGCTGAGGTGAAGGCAAGCAAGAGGAGAGGAGTGTGACAGCCAG 1661

Qy 1662 AGACTGCTCTCTCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 1721  
Db 1662 AGACTGCTCTCTCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 1721

Qy 1722 CCCACTGATCTGTGACAGAGGCTGCTGCT-CTGTTGCCAGGCTCTGTTTATAGCCATGAT 1780  
Db 1722 CCCACTGATCTGTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1780

Qy 1780 CCCACTGATCTGTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1783  
Db 1780 CCCACTGATCTGTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1783

QY	1781	CAGATGTGGTACGACTCTTTTCGCGGCGTGGAGACCAACGAGTCTGATGACGTCTCTGT	1840
Db	1784	CAGATGTGGTACGACTCTTTTCGCGGCGTGGAGACCAACGAGTCTGATGACGTCTCTGT	1843
QY	1841	GGACACGAGTGGTATGAGGATCTCAGGCGAGCCTCAGACCCCAACCTTCTACCTGCTTTGAC	1900
Db	1844	GGACACGAGTGGTATGAGGATCTCAGGCGAGCCTCAGACCCCAACCTTCTACCTGCTTTGAC	1903
QY	1901	TTGCTTTCTA-GCATAGCCTGGGCGCAAGCAGGCTGGGGAAATGAGAGATAGACATGGAGATG	1959
Db	1904	TTGCTTTCTAGCATAGCCTGGGCGCAAGCAGGCTGGGGAAATGAGAGATAG-CATGGAGATGT	1962
QY	1960	ATGGAGAGGATGGAGATTTTCCGCAAAAAAAAAAAAAAAAAAAAAA 2007	
Db	1963	ATGGAGAGGATGGAGATTTTCTATGTATAAATTAATAAAAAAAAAAAAA 2010	
RESULT 5			
LOCUS	BD006990	2065 bp	DNA linear PAT 31-JAN-2007
DEFINITION	Modulators of BRCA1 activity.		
ACCESSION	BD006990		
VERSION	BD006990.1	GI:18635361	
KEYWORDS	JP 2001502893-A/1.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 2065)		
AUTHORS	Luhinfield, B., Porakisu, P., Rigenferuta, C. and Buon, T.T.		
TITLE	Modulators of BRCA1 activity		
JOURNAL	Patent: JP 2001502893-A 1 06-MAR-2001;		
COMMENT	ONIX PHARMACEUTICALS INC		
	OS Unidentified		
	PN JP 2001502893-A/1		
	PD 06-MAR-2001		
	PF 06-AUG-1997 JP 1998512659		
	PR 04-SEP-1996 US 60/025601		
	PI BONY LUBINFELD, PAUL PORAKISU, CAROL RIGENFERUTA, TERIRIN T BUON		
	PC C12N15/09, A61K38/00, A61P43/00, C07K14/435, C07K16/18, C12N1/15,		
	PC C12N1/19,		
	PC C12N1/21, C12N5/10, C12P21/02, C12P21/08, C12N15/00, C12N5/00, PC		
FEATURES	A61K37/02		
	CC Strandedness: Double;		
	CC Topology: Linear;		
	FH Key Location/Qualifiers		
SOURCE	FT CDS Location/Qualifiers		
	1..2065		
BASE COUNT	561 a	526 c	561 g 417 t
ORIGIN			
Query Match 94.2%; Score 1890.8; DB 6; Length 2065;			
Best Local Similarity 98.9%; Pred. No. 0;			
Matches 1946; Conservative 0; Mismatches 17; Indels 5; Gaps 4			
QY	44	TACGAAGCCGGACCTGTACAGTTCCTTTGGCTGCGTGGGCCCTTTGAGTCCAGCCATCA	103
Db	44	TACGAAGCCGGACCTGTACAGTTCCTTTGGCTGCGTGGGCCCTTTGAGTCCAGCCATCA	103
QY	104	TGCGTATCCGTCGTCGTGACACTATCTGCTCGACTTTCGATGATCACTCCCGAGAGTGG	163
Db	104	TGCGTATCCGTCGTCGTGACACTATCTGCTCGACTTTCGATGATCACTCCCGAGAGTGG	163
QY	164	CCGCCATCCACTGCGGCCACACCTTCCACTTCCAGTGCCTTAATTCAGTCTTTGAGACAG	223
Db	164	CGGCATCCACTGCGGCCACACCTTCCACTTCCAGTGCCTTAATTCAGTCTTTGAGACAG	223
QY	224	CACCAAGTGGACCTGCGCCACAGTGGCGGAATCCAGGTGTGGCAAAAGACATTATCAATA	283
Db	224	CACCAAGTGGACCTGCGCCACAGTGGCGGAATCCAGGTGTGGCAAAAGACATTATCAATA	283

QY	284	AGCTCTTCCTTTGATCTTGGCCAGGAGAGAGATGCTTGGATCGAATATCTTTAAAGA	343
Db	284	AGCTCTTCCTTTGATCTTGGCCAGGAGAGAGAGATGCTTGGATCGAATATCTTTAAAGA	343
QY	344	ATGAACTGGACAAATGTCAAGAGCCCACTTTCCAGAAACAAGAAACGACACACC	403
Db	344	ATGAACTGGACAAATGTCAAGAGCCCACTTTCCAGAAACAAGAAACGACACACC	403
QY	404	AGGTCATCATGACACTCTCGGGGATACGCTGGAAACAAGCAATCTCTCTGGATATCTC	463
Db	404	AGGTCATCATGACACTCTCGGGGATACGCTGGAAACAAGCAATCTCTCTGGATATCTC	463
QY	464	TGCACAGGCGCTTGGGCAAGGCCGAGATGCTGTCTCCACTGTAAAAAGCAGATGAAGT	523
Db	464	TGCACAGGCGCTTGGGCAAGGCCGAGATGCTGTCTCCACTGTAAAAAGCAGATGAAGT	523
QY	524	ACTTAGAGACAGACAGAGATGAGAACCAACAACAGCACAGAGAGAGCGGGGCTCAGGA	583
Db	524	ACTTAGAGACAGACAGAGATGAGAACCAACAACAAGAGAGAGCGGGGCTCAGGA	583
QY	584	GCAAGATGAAGACCATGAGACAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGG	643
Db	584	GCAAGATGAAGACCATGAGACAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGG	643
QY	644	AGGAGATGATCCGAGACATGGGTGTGGGAGAGTCAGACGGGTGGACAGCTGGCTGTACT	703
Db	644	AGGAGATGATCCGAGACATGGGTGTGGGAGAGTCAGACGGGTGGACAGCTGGCTGTACT	703
QY	704	GTGTGTCTCTCAGAAAGAGTACGAGATCTAAAGAAGCCACGGAAGGCTTCAGGGAGG	763
Db	704	GTGTGTCTCTCAGAAAGAGTACGAGATCTAAAGAAGCCACGGAAGGCTTCAGGGAGG	763
QY	764	TGGCGTCAACCTGAGGAAGATTTGTTTCCCTCCGAAGCAAGTTGACGACAGTCTACT	823
Db	764	TGGCGTCAACCTGAGGAAGATTTGTTTCCCTCCGAAGCAAGTTGACGACAGTCTACT	823
QY	824	CTGAATTGGATCAGAGCCCAAGTTAGAACTGAAAGTCACACCCAGAAAGACTTACAGATGCTG	883
Db	824	CTGAATTGGATCAGAGCCCAAGTTAGAACTGAAAGTCACACCCAGAAAGACTTACAGATGCTG	883
QY	884	ACAAAGAAATCATGAGCTGAAAAAGAGCTAACGATGCTGCAGGAAACCTTGAACCTGC	943
Db	884	ACAAAGAAATCATGAGCTGAAAAAGAGCTAACGATGCTGCAGGAAACCTTGAACCTGC	943
QY	944	CACCAATGGCCAGTGAACATGTGACCGCGCTGGTTTAAAGAGCCAGCGCTGTGGAGG	1003
Db	944	CACCAATGGCCAGTGAACATGTGACCGCGCTGGTTTAAAGAGCCAGCGCTGTGGAGG	1003
QY	1004	TGAATCTGAACCTCGCGCGGCAATCCCTCCGTATATATGATCTCAATGCTACTCTTG	1063
Db	1004	TGAATCTGAACCTCGCGCGGCAATCCCTCCGTATATATGATCTCAATGCTACTCTTG	1063
QY	1064	ATGTGGATACTCCCCAGAGCCCGGCGCTCCAGCTCCAGCATGGTTACTACGAAAAACTTT	1123
Db	1064	ATGTGGATACTCCCCAGAGCCCGGCGCTCCAGCTCCAGCATGGTTACTACGAAAAACTTT	1123
QY	1124	GCCTTAGAGATGCACACTCCCCCAATTCAGATGTCCCCCAAGAAATATGCAAAAGCCCCCA	1183
Db	1124	GCCTTAGAGATGCACACTCCCCCAATTCAGATGTCCCCCAAGAAATATGCAAAAGCCCCCA	1183
QY	1184	GGAAGAGATCCCACTCTCACTGGGTGAGCCAGAGCTGTGCAAGAGAGCCAGATGAGAGAC	1243
Db	1184	GGAAGAGATCCCACTCTCACTGGGTGAGCCAGAGCTGTGCAAGAGAGCCAGATGAGAGAC	1243
QY	1244	TGTTGTGATCTTCCTATTTTGTCCGGAATCCATCTCTAGGCCAGAAACAGCCCAAA	1303
Db	1244	TGTTGTGATCTTCCTATTTTGTCCGGAATCCATCTCTAGGCCAGAAACAGCCCAAA	1303
QY	1304	GGCCCAAGTCAAGATCTCTTTGCAAGCAAGATGTGTAAGAGACAGGCTTTCGATGGGCTCG	1363
Db	1304	GGCCCAAGTCAAGATCTCTTTGCAAGCAAGATGTGTAAGAGACAGGCTTTCGATGGGCTCG	1363



OY	1364	GGGCGCGGCAAAATTCATCCAGCCCTACTGACACAGTCAATGTATCCGCCCATTTGCCCTGTTA	1423
OY	1364	GTTGGCGGCACAATAATTCACTCAGCCCTACTGAACACAGTCAATGTATCCGCCCATTTGCCCTGTTA	1423
OY	1424	AGCCCCAAGCACCAAGGTTTTAAGCAGAGGCTGAGGGTGAAGAAGCCCTGGCTTCTCTCTCCAGG	1483
Db	1424	AGCCCCAAGCACCAAGGTTTTAAGCAGAGGCTGAGGGTGAAGAAGCCCTGGCTTCTCTCTCCAGG	1483
OY	1484	CCAAGCTGACACCTTTCCTGCTGCTGTGAGAACAGTAGTGAGCTTGACCMAATGGCCAGACACA	1543
Db	1484	CCAAGCTGACACCTTTCCTGCTGCTGTGAGAACAGTAGTGAGCTTGACCMAATGGCCAGACACA	1543
OY	1544	TGCCGTGCAACCTGTAGGTGAAGGACGTGTCAGAGCAGG - TTTGGGAGAGAGCCCTACT	1601
Db	1544	TGCCGTGCAACCTGTAGGTGAAGGACGTGTCAGAGCAGG - TTTGGGAGAGAGCCCTACT	1603
OY	1602	TTTCGGGACCAGCCTGAGGTGTAAAGGGACAGAACACAGGTGAGGGTGTAGTATACACCAG	1661
Db	1604	TTTCGGGACCAGCCTGAGGTGTAAAGGGACAGAACACAGGTGAGGGTGTAGTATACACCAG	1663
OY	1662	AGACTGCTCTTCCTGSCCTCACCCCTGCCCACCTCTACAGACTGGAGGTGACATGACACG	1721
Db	1664	AGACTGCTCTTCCTGSCCTCACCCCTGCCCACCTCTACAGACTGGAGGTGACATGACACG	1723
OY	1722	CCCACTGATCCCTGTGACAGAGTCCTGCT - CTGTTGCCAGAGCTCTGTTATAGCCATGAT	1780
Db	1724	CCCACTGATCCCTGTGACAGAGTCCTGCT - CTGTTGCCAGAGCTCTGTTATAGCCATGAT	1783
OY	1781	CAGATGTGGTCAGACTCTTTCTGGGCGCTGAGAACACAGGTCACCTTGTTGACTGCTGCT	1840
Db	1784	CAGATGTGGTCAGACTCTTTCTGGGCGCTGAGAACACAGGTCACCTTGTTGACTGCTGCT	1843
OY	1841	GGACCAAGATGCTTGAGGCATCTACAGGACGCTTCAGCCCAACCTTCTACCTGCCCTTTCGAC	1900
Db	1844	GGACCAAGATGCTTGAGGCATCTACAGGACGCTTCAGCCCAACCTTCTACCTGCCCTTTCGAC	1903
OY	1901	TTTGCTTTCA - GCATAGCCCTGGGCGCAAGCAGGTTGGGGAATGCAGGATAGACATGGGAGCT	1959
Db	1904	TTTGCTTTCA - GCATAGCCCTGGGCGCAAGCAGGTTGGGGAATGCAGGATAGACATGGGAGCT	1962
OY	1960	ATGAGAGAGATGGAAGATTTTCCGAAAAAAAATTTAAAAAAAAAAAAA 2007	
Db	1963	ATGAGAGAGATGGAAGATTTTTCATGTTAAATTTAAATAAAAAA 2010	
RESULT	6		
LOCUS	BC017374	2024 bp mRNA linear ROD 14-NOV-2001	
DEFINITION	Mus musculus; TRAF-interacting protein, clone MGC:13919		
ACCESSION	BC017374		
VERSION	BC017374.1 GI:16924208		
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 2024)		
JOURNAL	Straussberg,R. Direct Submission Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org		

Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAK Plate: 18 Row: d Column: 5

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 6755868.

```

FEATURES
    source
        1..2024
            Location/Qualifiers
                /organism="Mus musculus"
                /db_xref="LocusID:22036"
                /db_xref="taxon:10090"
                /clone="MCC:13919 IMAGE:3990371"
                /tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMV."
                /clone_lib="NCI CGAP_Lu29"
                /lab_host="DH10B"
                /note="Vector: pcMV-SPORT6"
                /codon_start=1
                /product="TRAF-interacting protein"
                /protein_id="AAH17374.1"
                /db_xref="GI:16924209"
                /translation="MPRALCTTICSDFFDHSRDVAIHCGHTPHLOCLIQFETAPSRTCPQRIOVGKKTIINKLFDAEENVDLAELKNELDSVKNOISQDKREKDSQAIIIDLRLERNAIVTESLONALNKAMLTSLTKOKKPLEORODEFKOAREEHRLKCKIMMTDEIOELLLOSOREVEEMLRDMGVGSAYEOLAVYCVSLRKEYENLKERRATGEILDRLKKDLVSRSKLKTINTELDOKLELRSAQKDLOSADOMEISLKKMLLOGTLSPATNETVSRVLFPESPAVPEMMMRPLHPFGEDITFDVNPFTPTQTSQSHCLPRKLCLEARNPOMONLVLRKHVKSPPSOALSQROCVGELEDLAEPFLPIRNAAVLQKQPNRTTAASRCSTDVYRIGFGLGGRTFIOPRDTTIRPVVSKASKSKOKVRIKTVSSASPDKDFLFCQ"
BASE COUNT      582 a      489 c      528 g      425 t
ORIGIN

```

Query Match	Best Local Similarity	Score	DB 10;	Length	2024;
Matches 1492;	Conservative	0;	Mismatches 440;	Indels	Gaps
OY	26	AGCGAAGCGGGGCCCTCTACGAAGCCGGACCTGTAGCAGTTTCCTTGCGCTGGGCC	85		
Db	26	AGCGGCGCGCGTTTCGGCGACCACCAACTGTGTGTCGCTG-CGACCTGGTCCCGGGCT	84		
OY	86	CCTTAGTCCAGCCATCATMGCTATCCGTCGTGTGACTATCTGTCGCAGCTTCTTG	145		
Db	85	GCTTAGTCCAGCCATCATMGCTATCCGTCGTGTGACTATCTGTCGCAGCTTCTTG	144		
OY	146	ATCACTCCCGGACGTGGCGCCGACATCCACTGCGGCGCACACCTTCCACTTGGACGTCC	205		
Db	145	ATCACTCCCGGACGTGGCGCTGCGCATCCACTGTGGCGCACACTTTCATCTGCATAAGCC	204		
OY	206	TTCAGTCCCTTGGAGACAGACCACTGGACCTGCCACAAGTCCCGAATCCAGATTGGCA	265		
Db	205	TCCAGTGGTTGGAGACACACCAAGTCCGACCTGCCACAGTGTAGAATCCAGATTGGCA	264		
OY	266	AAAGAACATATCAATTAAGCTCTTGTGANTGTGGCCAGAGAGAGAGAAATGCTTGG	325		
Db	265	AAAAGACATATTAATAACAACCTTCTTGTGACCTTCGCCAGAGAGAGAAATGCTTGG	324		
OY	326	ATCGAATTTCTTAAAGAAATGAATGACATGTACAGAGCCAGCTTTCAGGAAGACA	385		
Db	325	ATCGAATTTCTTAAAGAAATGAATGACATGTACAGAGCCAGCTTTCAGGAAGACA	384		
OY	386	AGGAGAAACGAGACAGCGAGTCAATCGACACTCTGCGGGAATACGTTGAAGAAGCA	445		
Db	385	GGGAGAAACGAGACAGCGAGTCAATCGACACTCTACGAGGACACCTGGAAGAAGCA	444		
OY	446	ATGCTACTGTGTATCTCTGACAGAGCCCTTGGGCAAGGCCGAGATGCTGTCTCCACAC	505		
Db	445	ATGCTACTGTGTGTATCTCTGACAGAGCCCTTAAACAAAGCGAGAGATGCTGTCTCCAC	504		



[illegible][illegible]

TLSPRTNETVSRLEFSPAPVEMNRLHPPEDEIDINTTDEVTNPTOTSGSQ  
HLPKLELRARSPMVKLVKHKVSPESOLSGRORCEVDEELAGAEPLIRN  
AVLGKOPNRTTAESSRSTDYRIEFGDFDGLGGRTRKIQORDTTIIRPVPKSKAKSKQ  
VRKTVSASQPKDIFLQO"  
BASE COUNT 532 a 489 c 534 g 426 t  
ORIGIN

Query Match 53.9%: Score 1081; DB 10; Length 1981;  
Best Local Similarity 74.6%: Pred. No. 1.6e-291;  
Matches 1504; Conservative 0; Mismatches 450; Indels 63; Gaps 9;

1 GTGCGGTGAGCGAATTTGAAGCAAGCGAGCGGGGC--GCTTACGAAGCGCGACCT 58  
Db GTGCGGTGAGCGAATTTGAAGCAAGCGAGCGGGGTGCGGTCCACCAACTGTCT 74  
59 GTAGCAATTTCTTTGGCTGCTGGCCCTTGATGCACCATCATGCTATCGTGTTC 118  
Db GTCTCTG--GCAGCTGGTTCCTGGGCTGCTGTGATCGACCATCATGCTATCTCTTC 133  
119 TGTGCACTATGCTTCGCACTCTTGATCACCCTCGGACAGTGGCGGCATCCACTGG 178  
Db TGTGCACTATGCTTCGCACTCTTGATCACCCTCGGACAGTGGCGCATCCACTGTG 193  
179 GGCACACTTCCACTTGCACTGCTCAATTCAGTCTTTGAGACAGCACCAAGTGGACCT 238  
Db GGCACACTTTCATCTGCAATGCTCAATTCAGTGTTTAGACAGCACCAAGTGGACCT 253  
239 GGCACAGTGGCGAATCCAGGTTGGCAAAAGCACTTTCATTAAGCTCTCTTTGATC 298  
Db GGCACAGTGTAGAAATCCAGGTTGGCAAAAGCACTTTCATTAAGCAACTTTCTTTGAC 313  
299 TTTGCCAGAGAGAGAAATGCTTGATCGAATTTCTTAAGATGAAGTGGACATG 358  
Db TTTGCCAGAGAGAGAAATGCTTGATCGAATTTCTTAAGATGAAGTGGACATG 373  
359 TCGAGAGCCAGCTTTCGCCAAGAAAGCAAGAAAGCAGACAGCCAGTCAATCATGACA 418  
Db TCGAAGCTCAGCTTTCGCCAAGAAAGCAAGAAAGCAGACAGCCAGTCAATCATGACA 433  
419 CTCTGCGGATACGCTGGAAAGCAAGATGCTACTGTGTATCTCTGACAGAGCCTTGG 478  
Db CTCTACGGGACACCTGGAAAGCAAGATGCTACTGTGTATCTCTGACAGAGCCTTAA 493  
479 GCAAGCCAGATGCTGTCTCCACACTGAATAAGCAGATGAAGTGAAGTGAAGCAGCAG 538  
Db ACAAGGACAGATGCTGTCTCCACCTGAATAAGCAGATGAAGTGAAGTGAAGCAGCAG 553  
539 AGGATGAGACCAAGCAAGCAGAGAGAGCGGCGCGCTCAGAGCAAGATGACAGCA 598  
Db AGGATGAGACCAAGCAAGCAGAGAGAGCGGCGCGCTCAGAGCAAGATGACAGCA 613  
599 TGGAGCAGATTAAGCTTCTACTCCAGAGCAGCTCCCTGAGTGGAGAGATGATCCAG 658  
Db TGGACCAAAATTAAGCTTCTACTCCAGAGCAGCTCTTCTGAGTGGAGAGATGATCCAG 673  
659 ACATGGGTGGAGCAGTCAAGCGGAGCAAGCAGTGGCTGATGCTGTCTCTCAAA 718  
Db ACATGGGTGGAGCAGTCAAGCGGAGCAAGCAGTGGCTGATGCTGTCTCTCAAA 733  
719 AAGATGACAGATCTAAAGAGGACAGCAAGCGCTCAGAGGAGGTGCTGCAAGCTGA 778  
Db AAGATGACAGATCTAAAGAGGAGCGGAGGACAGCAAGCGGAGGTGCTGCAAGCTGA 793  
779 GGAAGATTTTCTTCTCCAGAGCAAGTGCAGACAGTCTACTCTGATTTGATGATCAG 838  
Db AGAAGATTTTCTTCTCCAGAGCAAGTGCAGACAGTCTCAACTGTGATGATCAG 853  
839 CCAATTTGAATGAAGTCAAGCCAGAGAGATTAAGAGTGTCTGCAAGAAATCATGA 898  
Db CCAATTTGAATGAAGTCAAGCCAGAGAGATTAAGAGTGTCTGCAAGAAATCATGA 913  
899 GCCTGAAAAAGCTAACGATGCTGCAGAGAAACCTTGAACCTGACACAGTGGCGAGTG 958

914 GCCTAAGAAAGAGCTGATGATCTCCAGAGAACTTGAGCCTGCGCTCCGCTACCAATG 973  
959 AAGACTGCAAGCCCTGCTGTTTGAAGAGCCAGCCCTGTGCA---GGTAATCTGAAGC 1015  
Db AAGAGGTCAAGCCCTGCTGTTTGAAGAGCCAGCCCTGTGATGATGAAGAACCCGAGGC 1033  
1016 TCCGCGGCGCATCTCCGTGATGATGATGATCTCAATGCTCACTTGTGATGATCTC 1075  
Db TCCGCGGCGCATCTCCGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093  
1076 CCCAGCCCGGCGCTCCAGCTCCAGCATGTTTACTAGCAAAACTTTGCTTAGAAGT 1135  
Db CTCCAACCCAGACCTCTGCTGCCAGCATGCTCCCAAGAAAGCTGCGCTGAGAGGG 1153  
1136 CACACTCCCAATTAAGATGTCCTCCAGAGATATGCAAGAGCCCGAGAGAGTCCC 1195  
Db CAGCTCTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1213  
1196 AGCTGCACTGGGTGGCGAGCTGTGAGAGAGCCAGATGAGAGTGGTGGTGGCT 1255  
Db AGCTGCACTGGGTGGCGAGCTGTGAGAGAGCCAGATGAGAGTGGTGGTGGCT 1273  
1256 TCCCTAATTTTCTCCGGAATGCCATCTAGGCCAGAAAGCCCAAAAGGCCAGTGA 1315  
Db TCCCTAATTTTCTCCGGAATGCCATCTAGGCCAGAAAGCCCAAAAGGCCAGTGA 1333  
1316 AGTCTCTTGACAGCAAGATGCTGTAAGACAGCTTGCATGGCTCGGTGGCCGACAA 1375  
Db AATCCGGAAGACAGCAAGATGCTGTAAGAGATGATGATGATGATGATGATGATGATGAT 1393  
1376 AATTCATCCAGCTACTGACAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1435  
Db AATTCATCCAGCTACTGACAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1453  
1436 AGTTAAGCAGAGGTGAGGTGAGAGCCGCTCTCTCTTCCAGGCGCAAGCTGACA 1495  
Db AGTTAAGCAGAGGTGAGGTGAGAGCCGCTCTCTCTCTTCCAGGCGCAAGCTGACA 1513  
1496 CCTTCTGTGCTGCTGAGAGACAGTGTGACCAATGGCCAGACATGCTGCACTT 1555  
Db CCTTCTGTGCTGCTGAGAGACAGTGTGACCAATGGCCAGACATGCTGCACTT 1588  
1556 GTAGGCAAGAGCTGACCA--GGCAGGGTGTGAGCAGAGCCCTACTTTGGGAGCAGC 1613  
Db GTAGGCAAGAGCTGACCA--GGCAGGGTGTGAGCAGAGCCCTACTTTGGGAGCAGC 1646  
1559 GTGGGCAAGAGCTGCTTACCGGAGAGTGTGAGAGATGCTCTCTTGGAGC---- 1614  
1614 CTGAGGTGTAAGGCGAGCAAAACAGGTGAGGTGAGTGCACCCAGAGAGCTGCTTTC 1673  
Db CTGAGGTGTAAGGCGAGCAAAACAGGTGAGGTGAGTGCACCCAGAGAGCTGCTTTC 1696  
1674 CTGCGCTCAAGCTGCGCCCACTCTTACGACTGGAGACTGACATGACAGCCCACTATCT 1733  
Db CTGCGCTCAAGCTGCGCCCACTCTTACGACTGGAGACTGACATGACAGCCCACTATCT 1756  
1647 CTGTTTCACTGCGCCCTGACCAAC--ACTGGAGAGCCACATGACAGTCACTTTCGG 1704  
1734 GTACAGAGTCTGCG--TGTGTTGAGAGGCTGTTTATGAGCATGATGATGAGTGA 1792  
Db GTACAGAGTCTGCG--TGTGTTGAGAGGCTGTTTATGAGCATGATGATGAGTGA 1815  
1705 ATCAGAGGAGCTTACTTCCAGTTCAGAGGTTTGTCTTATGCTTACACCAAGTGTGGCTG 1764  
1793 GACTCTTCTGCGGCTGAGACAGCAGTCACTTGTGATGCTGTCTGTGAGCAAG-- 1847  
Db GACTCTTCTGCGGCTGAGACAGCAGTCACTTGTGATGCTGTCTGTGAGCAAG-- 1870  
1765 GACTCTTCTTGTGTTTATGAGAAACAGGTCACATTCAGCTTAAGTGAAGTGGAGTCTGGA 1824  
1848 AGTCTTGAAGCATCTCAGAGAGCTCAGAGCCAGCTTCTACTGCTTGTGATGCTTTC 1907  
Db AGTCTTGAAGCATCTCAGAGAGCTCAGAGCCAGCTTCTACTGCTTGTGATGCTTTC 1930  
1825 GATCTCTATGAGAGCTGAGAGCCCTGCGCTTGAATCTGCTGCTCCAGCTTATGTC 1884  
1908 TAGCATATCGTGGCGCAAGCAGGCTGGGGAATGAGAGTATGACATGAGGATGATGAGAG 1967  
Db TTAGAATTAAGGAGTGAAGTGTGATGAGGAAAGGTTGGGAAAGTCTTGTGTAATAATTA 1944  
1968 GATGAGATTTTCCGAAAAAAGAAAAAAGAAAAA 2004



annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>

465N24 is from the library RPC13 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR>. pcrPAC.

## FEATURES

```
Source
1.129747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p35.1-36.13"
/clone="RP3-465N24"
/clone_lib="RPC1-3"
540..753
repeat_region
/note="L2 repeat: matches 2255..2461 of consensus"
975..1303
repeat_region
/note="AluJb repeat: matches 1..311 of consensus"
1320..1547
repeat_region
/note="MIR repeat: matches 29..262 of consensus"
2061..2268
repeat_region
/note="MIR repeat: matches 100..262 of consensus"
2356..2518
repeat_region
/note="MERSA repeat: matches 5..164 of consensus"
2519..2827
repeat_region
/note="AluJb8 repeat: matches 1..318 of consensus"
3245..3360
repeat_region
/note="MIR repeat: matches 20..139 of consensus"
3427..3485
repeat_region
/note="L2 repeat: matches 2688..2746 of consensus"
3444..3676
repeat_region
/note="MIR repeat: matches 8..249 of consensus"
3690..3856
repeat_region
/note="LTR41 repeat: matches 20..182 of consensus"
4187..4494
repeat_region
/note="AluSg repeat: matches 1..308 of consensus"
4846..5143
repeat_region
/note="AluSx repeat: matches 1..296 of consensus"
5287..5340
repeat_region
/note="L1MD1 repeat: matches 5345..5398 of consensus"
5361..5518
repeat_region
/note="L1MA4 repeat: matches 6152..6300 of consensus"
5508..6164
repeat_region
/note="L1MD1 repeat: matches 5423..6091 of consensus"
6176..6245
repeat_region
/note="35 copies 2 mer ag 67% conserved"
6247..6528
repeat_region
/note="AluSg repeat: matches 15..291 of consensus"
6529..6674
repeat_region
/note="73 copies 2 mer ta 73% conserved"
6741..6874
repeat_region
/note="L1MD1 repeat: matches 6090..6223 of consensus"
7225..7278
repeat_region
/note="MIR repeat: matches 36..85 of consensus"
8137..8448
repeat_region
/note="AluSx repeat: matches 1..312 of consensus"
9092..9137
repeat_region
/note="23 copies 2 mer at 100% conserved"
9229..9326
repeat_region
/note="L2 repeat: matches 2574..2670 of consensus"
9669..9732
repeat_region
/note="MIR repeat: matches 87..148 of consensus"
10162..10365
repeat_region
/note="MIR repeat: matches 8..222 of consensus"
10379..10505
repeat_region
/note="L2 repeat: matches 2586..2710 of consensus"
10654..10814
repeat_region
/note="MERS6 repeat: matches 9..176 of consensus"
11454..11752
repeat_region
/note="AluJb repeat: matches 1..302 of consensus"
11790..12171
repeat_region
/note="L1M4 repeat: matches 1258..1640 of consensus"
12172..12362
repeat_region
/note="MERS3 repeat: matches 1..189 of consensus"
12363..12775
repeat_region
/note="L1M4 repeat: matches 1640..2049 of consensus"
12833..13130
repeat_region
/note="AluJb repeat: matches 1..291 of consensus"
13476..14035
repeat_region
/note="L1M4 repeat: matches 2224..2742 of consensus"
14036..14330
repeat_region
/note="AluJ repeat: matches 1..301 of consensus"
14331..14466
repeat_region
/note="L1M4 repeat: matches 2742..2867 of consensus"
14467..14600
repeat_region
/note="FLAMC repeat: matches 1..133 of consensus"
14601..14885
repeat_region
/note="L1M4 repeat: matches 2867..3278 of consensus"
14893..15096
repeat_region
/note="MTRD repeat: matches 286..505 of consensus"
15159..15484
repeat_region
/note="MTRD repeat: matches 6..323 of consensus"
15487..16041
repeat_region
/note="L1MC4 repeat: matches 7451..7973 of consensus"
16069..16773
repeat_region
/note="L1ME3A repeat: matches 3211..3943 of consensus"
16774..17068
repeat_region
/note="AluSg repeat: matches 3..297 of consensus"
17071..17384
repeat_region
/note="AluSg repeat: matches 1..310 of consensus"
17385..17451
repeat_region
/note="L1ME3A repeat: matches 3943..3955 of consensus"
17567..17775
repeat_region
/note="L1ME3A repeat: matches 4012..4208 of consensus"
17912..18219
repeat_region
/note="AluSg repeat: matches 1..313 of consensus"
18401..18922
repeat_region
/note="L1ME3A repeat: matches 4315..4848 of consensus"
18923..19226
repeat_region
/note="AluSx repeat: matches 1..304 of consensus"
19227..19530
repeat_region
/note="L1ME3A repeat: matches 4848..4586 of consensus"
19531..19897
repeat_region
/note="L1PB1 repeat: matches 5781..6155 of consensus"
19898..20698
repeat_region
/note="L1ME3A repeat: matches 4585..5403 of consensus"
21061..21150
repeat_region
/note="MIR repeat: matches 115..213 of consensus"
21218..21390
repeat_region
/note="L2 repeat: matches 1626..1814 of consensus"
21511..21767
repeat_region
/note="L1MC4 repeat: matches 6664..6940 of consensus"
21771..21969
repeat_region
/note="L1MC2 repeat: matches 3850..4054 of consensus"
21970..22008
repeat_region
/note="u2 repeat: matches 1..39 of consensus"
22009..22354
repeat_region
/note="L1MC2 repeat: matches 4054..4410 of consensus"
22355..22663
repeat_region
/note="AluSx repeat: matches 1..309 of consensus"
22664..23540
repeat_region
/note="L1MC2 repeat: matches 4410..5261 of consensus"
23552..23730
repeat_region
/note="FAM repeat: matches -135..43 of consensus"
23731..24657
repeat_region
/note="L1MC2 repeat: matches 5278..6319 of consensus"
24659..24758
repeat_region
/note="AluJ/FLAM repeat: matches 2..77 of consensus"
24759..25029
repeat_region
/note="L1MC4 repeat: matches 7461..7777 of consensus"
25872..26333
repeat_region
/note="L1MB5 repeat: matches 4928..5368 of consensus"
```

```

repeat_region      26334..26622
                    /note="L1M48 repeat: matches 5995..6287 of consensus"
repeat_region      26623..26774
                    /note="L1M5 repeat: matches 5368..5512 of consensus"
repeat_region      26775..27077
                    /note="AluX repeat: matches 1..298 of consensus"
repeat_region      27078..27656
                    /note="L1M5 repeat: matches 5512..6176 of consensus"
repeat_region      27713..27944
                    /note="AluY repeat: matches 74..299 of consensus"
repeat_region      27959..28020
                    /note="L1M3 repeat: matches 6941..7003 of consensus"
repeat_region      28028..28269
                    /note="121 copies 2 mer gg 60% conserved"
misc_feature        28146..28221
                    /note="candem repeat in single clone region"
repeat_region      28512..28663
                    /note="L2 repeat: matches 2241..2422 of consensus"
repeat_region      29643..30184
                    /note="LRR2 repeat: matches 1..505 of consensus"
repeat_region      30368..30450
                    /note="MIR repeat: matches 28..118 of consensus"
repeat_region      31248..31409
                    /note="L2 repeat: matches 1794..1967 of consensus"
repeat_region      31416..31494
                    /note="MIR repeat: matches 187..261 of consensus"

```

Query Match	13.08;	Score 260;	DB 9;	Length 129747;
-------------	--------	------------	-------	----------------

Matches 547; Conservative 0; Mismatches 140; Indels 113; Gaps 9;

QY	1221	TCGACGAGACGACGATGAGGAACTGGTTGGGCCCTCCCTCAATTTTGTGCGGAATGCCAT	1280
Db	25030	TTCAATGGTGCTGATTTACTAAAGTATCGAATGCGTTCACTATTTTTCATTC -GAATGCCAT	25088
QY	1281	CCTAGGCCAGAAACAGCCCAAAAGGCCAGGTCAAGTCTCTTTCAGCCAAAGATGTGT	1340
Db	25089	CCTGGCTCGAAGAACAGCCCAAGAGGAGACAGACAGAGCCCTCTCCACCAAGGTGTGT	25148
QY	1341	AAGGACAGCGTTTCGATGGGCGTGGGGCCGGCAAAATTCATCCAGCCCTACTGACACAGT	1400
Db	25149	AAGGACA-----TCAATGGGCTTTGGTGGCTGGAAAAATTCATCCAGCCCTACTGACCCAGC	25204
QY	1401	CATGATCCGCCCATTTGCGTTTAAGCCCAAGACCAAGGTTAAGCAGAGGGTGAAGGTTGA	1460
Db	25205	CATGATCCGCCCATTTGCCATTAACTTAACTCCAAAGGCCAAGGTTAAGCAGAGAG-----	25254
QY	1461	GACCGTGCCCTTCTCTCTTCCAGGCCAAGCTGGACACCTTCCGTGGTGGTGAAGAACATG	1520
Db	25255	-----	25254
QY	1521	AGTGTGACCAATTGGCCAGACACATCCCTGCAACTGTGAGTTCAGAGACTGTCAGGACAG	1580
Db	25285	-----CAACCAATTCTCTTAACTGTGAGGCCAAGGGCTGGCTAAGGACAG	25299
QY	1581	G---TTTGTGCACAGAGCCCTTACTTTTCGGGACCAAGCTGAGSTGTAAAGGGCAGACAAACA	1637
Db	25300	GGGTTTAAAGGGGATTAACCCCTCTTGTAGGACACCTTGAAGAGTAAAGGGCAGACAAAGA	25359
QY	1638	GGTGAAGGTTGAGTGTGACACCCAGAGAGAC -TGCCTCTTCTGGCCCTCACCCCTCCACATC	1695
Db	25360	GGTGAAGGTTAAGTGTGACACACAGAGACTGTCTTTCTCGCCCTCACCCCTCCACATC	25419
QY	1696	CTACGACTGGGAGC-----TGCATGACCAAGCCCACTGATCTGTCAAGAGTCT	1746
Db	25420	CCACCAACATTGGGCTTACCCTTGGGAGGTGGGATTCCACGATCTCTCTGGCAGAGGCT	25479
QY	1747	GCT-CTGTTGCAGGCTCTTGTTTAAAGCATGATCAGATGTGGTGCAGACTCTTCTGTGG	1805
Db	25480	GCTACCTGTGCTAAGACTCTTATTTACAGCCATGATATAGACCAAGGCTCTTCTGTGT	25539
QY	1806	CGTGAGACCAAGCTCACTTGTGACTGTCTCTGTGAGACAGAGTGTGAGGATCTCA	1865

Db	25540	CTGAGAGCCAGAGTCATTTTGTGTACCTGTCTCTATGACCAAAAGTCTTTGAAGCATCTCA	25599
QY	1866	GGCAGCCTCAGGCCAAGCCTTCTACCTGCCTTTGACTGCTCTCTA-GCATAGCCTGGGCCA	1924
Db	25600	GGCAGACTTAACCTGAGTTTCTTACTCGCTCGACTTGCTGTATTAGCATGTCTCGGGCA	25659
QY	1925	AGCAGGTTGGGGAAATGAGAGTAGACATGGATGTATGGAGAGATGCAAGATTTTCCCG	1984
Db	25660	AGCAGGTTGGGGAAATGAGAGGAT-----GGAGTGTGTGAGAGGATGAATAATTTTCATG	25712
QY	1985	AAAAAAAAAAAAAAAAAAAAA	2004
Db	25713	TAAAACTTTTAAAAAGAA	25732

## RESULT 10

LOCUS	G23637	391 bp	DNA	linear	STS 31-MAY-1996
DEFINITION	human STS WI-13434, sequence tagged site.				

```

VERSION      G23637.1  GI:1343963
REVISIONDC   CMC      CMC      securedb     securedb     based     of the

```

**SOURCE** human STSS derived from sequences in dbEST and the Unigene collection

ORGANISM      Homo sapiens  
 Filikaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:

REFERENCE 1 (bases 1 to 391)  
Mammalia; Eulielia; Primates; Catarrhini; Hominoidea; Homo.

TITLE Whitehead Institute/MIT Center for Genome Research; Physically

JOURNAL Unpublished

**Contact: Thomas Hudson**

Whitehead Institute for Biomedical Research  
9 Cambridge Center Cambridge MA 02142 USA

Tel: 617 232 1300  
Fax: 617 252 1902

0  
1  
2  
3  
4  
5  
6  
7  
8  
9

Primer B: CTGTGACCAAGAGTCTTGA

PCR Profile:

Denaturation: 56 degrees C  
Annealing: 56 degrees C

PCR Cycles: 35

Protocol:

Primer: each 5 pm

Taq Polymerase: 0.025 units/ul

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

MgCl<sub>2</sub>: 1.5 mM  
KCl: 50 mM

1115 HOLD. 10 MIN  
PH: 9.3

Derived from dbEST (genbank accession R40580).

source	1. .391
--------	---------

```
/db_xref="taxon:9606"
```

Primer bind	STS
32	.181
33	.55

```
compress(100, :T8T)
primer_bind
```

BASE COUNT 87 a 93 c 106 g 99 t 6 others  
 ORIGIN

Query Match 10.6%; Score 212.8; DB 11; Length 391;  
 Best Local Similarity 92.6%; Pred. No. 2.8e-48;  
 Matches 289; Conservative 0; Mismatches 14; Indels 9; Gaps 6;

QY 1704 GGAGGTCAGCATGACC-ACCCAGCATGCTCTGACAGAGCTCTG---CTGTGTGGCCA 1758  
 DB 321 GGAAGCTGACATGACCCAGCCAGCCAGTGCCTGACAGAGCTCTGCTGTGGCCC 262  
 QY 1759 GGCTCTTGTATATAG-CCATGATCAGATGTGTGACACTTCTGTGGG-CCGTGAGACCA 1816  
 DB 261 GGCTCCGTTTATAGCCATGATGATGTGTGACACTTCTGTGGCCCTGAGACCA 202  
 QY 1817 CGGTCACTTGTACGTCTCTGTGACACAGAGTCTGTGAGCATCTCAGGACCTCAG 1876  
 DB 201 CGGTCACTTGTGTACGTCTCTGTGACACAGAGTCTGTGAGCATCTCAGGACCTCAG 142  
 QY 1877 CCCAGCTTCTACCTGCTTGTGACTTGTCTT-CTAGCATACCTGGGCGAAGAGGGTGGG 1935  
 DB 141 CCCAGCTTCTACCTGCTTGTGACTTGTCTTCTTCTAGCATACCTGGGCGAAGAGGGTGGG 82  
 QY 1936 GAATGAGAGATAGACATGGGATGTATGAGAGAGATGAGATTTTCCGAAAAA 1995  
 DB 81 GAATGAGAGATAG-CATGGATGTATGAGAGAGATGAGATTTTCAATTAATAAT 23  
 QY 1996 AAAAAAAAAA 2007  
 DB 22 TAAAAAAAAA 11

## RESULT 11

AX350344 AX350344 7542 bp DNA linear PAT 06-FEB-2002  
 LOCUS Sequence 3 from Patent WO0192492.  
 DEFINITION AX350344  
 ACCESSION AX350344  
 VERSION AX350344.1 GI:18616006  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 TITLE 1 (sites)  
 AUTHORS Yan,C.C., Wei,M.H., Ketchum,K.C., Merkulov,G.C. and Beasley,E.M.  
 JOURNAL Isolated human kinase proteins, nucleic acid molecules encoding  
 Patent: WO 0192492-A 3 06-DEC-2001;  
 Apiera Corporation Robert A. Millman Assistant Secretary (US)  
 FEATURES Location/Qualifiers

source 1..7542  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 1612 a 1977 c 2156 g 1797 t  
 ORIGIN

Query Match 8.8%; Score 176.4; DB 6; Length 7542;  
 Best Local Similarity 95.1%; Pred. No. 6.5e-38;  
 Matches 194; Conservative 0; Mismatches 6; Indels 4; Gaps 1;

QY 1 GTGCGTGGAGCGAAATTGAGCAAGCGGAGGGGGG---CTCTAGCAAGCCGGAGC 56  
 DB 7302 GTGCTCGGGGCGCAAAATTGAGCAAGCGGAGGGGGGCGGCTCTAGCAAGCCGGAGC 7361  
 QY 57 CTGTACAGATTCTTGGCTGCTGGGCCCTTGAGTCCAGCATATGCTTCCGTG 116  
 DB 7362 CTGTACAGATTCTTGGCTGCTGGGCCCTTGAGTCCAGCATATGCTTCCGTG 7421  
 QY 117 TCTGTCACTATCTGCTCGACTTCTTCATCATCTCCCGGAGAGTGGCGGCATTCAC 176  
 DB 7422 TCTGTCACTATCTGCTCGACTTCTTCATCATCTCCCGGAGAGTGGCGGCATTCAC 7481

QY 177 CGGCACACCTTCCACTTCAGTG 200  
 DB 7482 CGGCACACCTTCCACTTCAGTG 7505

## RESULT 12

AC099330 163424 bp DNA linear HTG 09-NOV-2001  
 LOCUS Homo sapiens chromosome 3 clone RP11-78010, WORKING DRAFT SEQUENCE,  
 DEFINITION 8 unordered pieces.  
 AC099330 AC068701  
 AC099330.1 GI:16874864  
 VERSION HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.  
 TITLE 1 (bases 1 to 163424)  
 JOURNAL Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.  
 REFERENCE Direct Submission  
 AUTHORS 2 (bases 1 to 163424)  
 JOURNAL Submitted (09-NOV-2001) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA  
 COMMENT On Nov 9, 2001 this sequence version replaced gi:8247914.

## COMMENT

Center: University of Washington Genome Center  
 Center Code: UMG  
 Web site: http://www.genome.washington.edu  
 Contact: uwgchgs@u.washington.edu  
 Drafting Center: BCM  
 ----- Project Information  
 Center project name: chr-3  
 Center clone name: RP11-78010 (bc0618)  
 ----- Summary Statistics  
 Sequencing vector: unknown; 42% of reads  
 Chemistry: Dye-terminator ET; 92% of reads  
 Chemistry: Dye-terminator Big Dye; 8% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 158969 bases at least Q40  
 Consensus quality: 160979 bases at least Q40  
 Consensus quality: 162023 bases at least Q20  
 Insert size: 202800; 21.9% error; agarose-tp  
 Insert size: 162724; sum-of-ctnigs  
 Quality coverage: 9.0x in Q20 bases; agarose-tp  
 Quality coverage: 11.2x in Q20 bases; sum-of-ctnigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 4608: contig of 4608 bp in length  
 \* 4609 4708: gap of unknown length  
 \* 4709 10289: contig of 5581 bp in length  
 \* 10290 10389: gap of unknown length  
 \* 10390 17105: contig of 6716 bp in length  
 \* 17106 17205: gap of unknown length  
 \* 17206 45100: contig of 27895 bp in length  
 \* 45101 45200: gap of unknown length  
 \* 45201 65174: contig of 19974 bp in length  
 \* 65175 65274: gap of unknown length  
 \* 65275 90719: contig of 25445 bp in length  
 \* 90720 90820: gap of unknown length  
 \* 90820 119377: contig of 28558 bp in length  
 \* 119378 119477: gap of unknown length

FEATURES \* 119478 163424: contig of 43947 bp in length.  
Location/Qualifiers  
source 1. 163424  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-78010"  
/clone\_1lb="RPci human BAC library 11"  
1. 4608  
misc\_feature /note="assembly\_name:Contig74"  
4709. 10289  
misc\_feature /note="assembly\_name:Contig75"  
10390. 17105  
misc\_feature /note="assembly\_name:Contig76"  
17206. 45100  
misc\_feature /note="assembly\_name:Contig77"  
45201. 65174  
misc\_feature /note="assembly\_name:Contig78"  
65275. 90719  
misc\_feature /note="assembly\_name:Contig79"  
90820. 119377  
misc\_feature /note="assembly\_name:Contig80"  
119478. 163424  
misc\_feature /note="assembly\_name:Contig81"  
BASE COUNT 40950 a 40937 c 39779 g 40912 t 846 others  
ORIGIN

Query Match 8.8%; Score 176.4; DB 2; Length 163424;  
Best Local Similarity 95.1%; Pred. No. 9.9e-38;  
Matches 194; Conservative 0; Mismatches 6; Indels 4; Gaps 1;

QY 1 GTGCGGTGAGCGAATTGTAACAGCGGAGCGGGGCGC---CTCTAGCAAGCCGCGAC 56  
Db 149275 GTGCTGGGAGCCAAATTGAACAGCGGAGCGGCGCGCTCTAGCAAGCCGCGAC 149334  
QY 57 CTGTACAGATTTCCTTGGCTGCTGGGCCCTTGGAGTCCAGCCATCATGCTATCCGTGC 116  
Db 149335 CTGTACAGATTTCCTTGGCTGCTGGGCCCTTGGAGTCCAGCCATCATGCTATCCGTGC 149394  
QY 117 TCTGTGACATATGCTCGACATCTTTCGATCACTCCCGGAGCGTGGCCGCAATCCACTG 176  
Db 149395 TCTGTGACATATGCTCGACATCTTTCGATCACTCCCGGAGCGTGGCCGCAATCCACTG 149454  
QY 177 CGGCGACACCTTCACATTCGACGTG 200  
Db 149455 CGGCGACACCTTCACATTCGACGTG 149478

RESULT 13  
AL353694\_2/c  
WPCOMMENT  
Sequence split into 4 fragments LOCUS AL353694 Accession AL353694  
Fragment Name Begin End  
AL353694\_0 1 110000  
AL353694\_1 100001 210000  
AL353694\_2 200001 310000  
AL353694\_3 300001 356330  
Continuation (3 of 4) of AL353694 from base 200001 (AL353694 Homo sapiens chromosome 6 c

Query Match 6.5%; Score 130; DB 2; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 9.4e-25;  
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 381 AGACAAGAGAAAGAGACAGCCAGGTATCATCATCTCTCGGAGATACGCTGAAGA 440  
Db 62746 AGACAAGAGAAAGAGACAGCCAGGTATCATCATCTCTCGGAGATACGCTGAAGA 62687  
QY 441 AGCAATGCTACTGTGTATCTCTGACAGCAGCCCTTGGGCAAGCCGAGATGCTGTGCTC 500  
Db 62686 AGCAATGCTACTGTGTATCTCTGACAGCAGCCCTTGGGCAAGCCGAGATGCTGTGCTC 62627

QY 501 CACACTGAAA 510  
Db 62626 CACACTGAAA 62617

RESULT 14  
AC099330 163424 bp DNA linear HTG 09-NOV-2001  
Homo sapiens chromosome 3 clone RP11-78010, WORKING DRAFT SEQUENCE,  
8 unordered pieces.  
AC099330 AC068701  
AC099330.1 GI:16874864  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULFILL; HTGS\_ACTIVEFIN.  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 163424)  
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
Direct Submission  
Unpublished  
2 (bases 1 to 163424)  
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
Submitted (09-NOV-2001) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Nov 9, 2001 this sequence version replaced gi:8247914.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Genome Center  
Center: University of Washington Genome Center  
Center Code: UWGC  
Web site: <http://www.genome.washington.edu>  
Contact: uwgchls@u.washington.edu  
Drafting Center: BCM  
----- Project Information  
Center project name: chr-3  
Center clone name: RP11-78010 (bc0618)  
----- Summary Statistics  
Sequencing vector: unknown; 42% of reads  
Chemistry: Dye-terminator ET; 92% of reads  
Chemistry: Dye-terminator Big Dye; 8% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 158969 bases at least Q40  
Consensus quality: 160979 bases at least Q30  
Consensus quality: 162023 bases at least Q20  
Insert size: 202800; 21.9% error; agarose-fp  
Insert size: 162724; sum-of-ctrls  
Quality coverage: 9.0x in Q20 bases; agarose-fp  
Quality coverage: 11.2x in Q20 bases; sum-of-ctrls  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 4608: contig of 4608 bp in length  
\* 4709: gap of unknown length  
\* 4709 10289: contig of 5581 bp in length  
\* 10290 10389: gap of unknown length  
\* 10390 17105: contig of 6716 bp in length  
\* 17106 17205: gap of unknown length  
\* 17206 45100: contig of 27895 bp in length  
\* 45101 45200: gap of unknown length  
\* 45201 65174: contig of 19974 bp in length  
\* 65175 65274: gap of unknown length  
\* 65275 90719: contig of 25445 bp in length  
\* 90720 90819: gap of unknown length  
\* 90820 119377: contig of 28558 bp in length  
\* 119378 119478: gap of unknown length  
\* 119478 163424: contig of 43947 bp in length.



FEATURES	SOURCE
misc_feature	location/Qualifiers 1..163424 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-78010" /clone.lib="RPCI human BAC library 11"
misc_feature	1..4608 /note="assembly_name:Contig74"
misc_feature	4709..10289 /note="assembly_name:Contig75"
misc_feature	10390..17105 /note="assembly_name:Contig76"
misc_feature	17206..45100 /note="assembly_name:Contig77"
misc_feature	45201..65174 /note="assembly_name:Contig78"
misc_feature	65275..90719 /note="assembly_name:Contig79"
misc_feature	90820..119377 /note="assembly_name:Contig80"
misc_feature	119478..163424 /note="assembly_name:Contig81"
BASE COUNT	40950 a 40937 c 39779 g 40912 t 846 others
ORIGIN	
Query Match	6.5%; Score 130; DB 2; Length 163424; Best Local Similarity 100.0%; Pred. No. 1e-24;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY 381 AGACAAGAGAAGACGACGACCAGGTATCATGCACCTCGGGGATACCTGGGAAGA 440	
Db 57007 AGACAAGAGAAGACGACGACCAGGTATCATGCACCTCGGGGATACCTGGGAAGA 56948	
OY 441 ACGCAATGCTACTGTCGTATCTGCACGAGCCCTTGGCAAGGCCGATGCTGTGCTC 500	
Db 56947 ACGCAATGCTACTGTCGTATCTGCACGAGCCCTTGGCAAGGCCGATGCTGTGCTC 56888	
OY 501 CACACTGAAA 510	
Db 56887 CACACTGAAA 56878	
RESULT 15	
LOCUS FRI165K09/c	39410 bp DNA linear VRT 23-JAN-2002
DEFINITION Fugu rubripes cosmid 165K09 DNA for GRM7, TRIF, Sand, PRGRF3 genes.	
ACCESSION AJ010317	
VERSION AJ010317.1 GI:3928163	
KEYWORDS GRM7 gene; metabotropic glutamate receptor 7; plasminogen related growth factor receptor 3; PRGRF3 gene; Sand gene; TRAF interacting protein; trip gene.	
SOURCE Takifugu rubripes.	
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Takifugu.	
REFERENCE 1 (sites)	
AUTHORS Cottage,A., Clark,M., Hawker,K., Umranta,Y., Wheller,D., Bishop,M. and Elgar,G.	
TITLE Three receptor genes for plasminogen related growth factors in the genome of the puffer fish Fugu rubripes	
JOURNAL PUBMED 99148833	
MEDLINE 10025966	
PUBMED 2 (sites)	
AUTHORS Cottage,A.J., Edwards,Y.J. and Elgar,G.	
TITLE SAND, a new protein family: from nucleic acid to protein structure and function prediction	
JOURNAL Compar. Funct. Genom. 2, 226-235 (2001)	
REFERENCE 3 (bases 1 to 39410)	
AUTHORS Cottage,A.J.	
TITLE Direct Submission	
JOURNAL Submitted (15-AUG-1998) Cottage A.J., HGMPRC, Computing, Hinxton, Cambridge, CB10 1SB, UK	
FEATURES	
source location/Qualifiers 1..39410 /organism="Takifugu rubripes" /db_xref="taxon:31033" /clone="cosmid 165K09" join(<261..401,496..696,1891..2030,2155..2778,2854..3165,4227..>4466) /gene="GRM7" 261..4466 /gene="GRM7" /join(<261..401,496..696,1891..2030,2155..2778,2854..3165,4227..4466) /gene="GRM7" /product="metabotropic glutamate receptor 7" /protein_id="CAA09083.1" /db_xref="GI:3928164" /db_xref="SPTRMBL:O9PMO0" /translation="FOQFTSRSLIENRRNIMFAEWEDEDFCKLTPIPKIKELGRK CTDERIDSDSOYEDEGKVOPVIDAYAAHAHLSHMHPTLCPSMGSCVKMPVSGRM LLQYIHGVNPNGSAGTGVPFNENGDAFGKRDIFQOMSNISPGYNICQMNHRLRN LEBMWNSGDNRKTIPESVCSFPCBSGERKKRVKGVPCCMHCELDGTYILDEFTDCMC PYDMRPLKNRGCRPPPIIKLEWSSPMALIPVFLAILGLIATGVATPIRENDPPIV RASGRELSVLLTLGLFIYLIFLMIAEPSVAVCARRFLIGMICISYSAMLTNR IYRIFEOGKSSTVPKPFISPTSQLITTFILISVOLLGVPFMGVMPPHIIIDEOKP PNEFPARGVLKCDMSDLTILISVSLIMTCTCYAIVKRSQVEPNFKAKIGTFMY TTCIVLAAPVFPIPTFGASTEKRFITOTTLTVSMISLTSAGMLTIIPVYVIITHPE ONYOKNRKRSFKAVAYONATVSTRLOKSSDKONGESKITEPDRSQ"	
exon number=1 /gene="GRM7" /number=1 402..495 /gene="GRM7" /number=1 496..696 /gene="GRM7" /number=2 697..1880 /gene="GRM7" /number=2 1891..2030 /gene="GRM7" /number=3 2031..2154 /gene="GRM7" /number=3 2155..2778 /gene="GRM7" /number=4 2779..2853 /gene="GRM7" /number=4 2854..3165 /gene="GRM7" /number=5 3166..4226 /gene="GRM7" /number=5 4227..>4466 /gene="GRM7" /number=6 complement(join(<4660..4801,6176..6226,6468..6521,6616..6743,6799..6931,7017..7106,7205..7292,7382..7495,7643..7737,7816..7943,8029..8068,8194..8277,8358..8415,8496..>8593)) /gene="TRIP" complement(4660..8593) /gene="TRIP" complement(join(4660..4801,6176..6226,6468..6521,	



[illegible]

LOCUS	AC067854	203984 bp	DNA	linear	HTG 04-JAN-2002
DEFINITION	Homo sapiens chromosome 8 clone RP11-513H8	map 8,	***	SEQUENCING IN	
PROGRESS	***	1 ordered	pieces.		

VERSION	ACC067854.13	GI:18056669
KEYWORDS	HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCES  
AUTHORS

1 (bases 1 to 203964)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 8, clone RP11-513H8  
Unpublished  
2 (bases 1 to 203984)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

TITLE	Direct Submission
JOURNAL	Submitted (27-Apr-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Jan 4, 2002 this sequence version replaced gi:18034761. All repeats were identified using RepeatMasker.

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information

```
Center project name: L9308
Center clone name: 513_H_8
-----
```

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* been provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* I 203984: contig of 203984 bp in length.

```
location/variables
1. .203984
/organism="Homo sapiens"
```

```

        /db_xref="taxon:9606"
        /chromosome="8"
        /map="8"
        /clone="RP11-513H8"
        /clone_lib="RPCT-11 Human Male BAC"
61400 a 39486 c 39964 g 63134 t

```

3.0%; Score 60.6; DB 2; Length 203984;  
1 Similarity 44.8%; Pred. No. 2.9e-05;  
280; Conservative 0; Mismatches 339; Indels 6; Gaps 1;

GGAGGAGGAGATGTCCTTGATTCGAGATTCTTAAGATGAACTGGACAAATGTCAGAG 364  
||||| | | ||| | | | | | | |  
||||| | | ||| | | | | | | | 14550

CCAGCTTCCAGAAAGACAAGAGAAACGAGACAGCCAGGTGTCATCATCGACACTCTGC 424

GGGAGCAGAGAAGGAGAGACGAGCAGAGAAGGAGAGACGAGCAAGAGAGCAGACGC  
CGCTTCCCCCTCCAAACAACCGAATTCCTTTACTGTCTGCTATTATCATCTGCCAACCC

184

[illegible]

CCGAGATGCTGTGTCACACTGAAAAAGCAGATGAAGTACTTAGAGCAGCAGCAGCATG 544  
||||| | ||||| ||||| ||||| |||||  
GAGCAAGGAG-----GAGGAGGAGAAGGAGGAGGAGAAGGAGGAGGAGGAGGAAA 16385

33ACCAACACAGCACACAGAGAGAGCGCGCGGCTCAGAGCAAGATGAGAGACCATTGAGC 604

GAATTGAGCTTCTACTCCAGAGCGAGCTCCCTGAGGTGAGGAGATGATCCGAGCATGG 664  
CGACGAGAGGAGGAGGAAAAAGCAGCAGCAGCAGGAGAGAGCAGCAGCAGG 16325

[illegible][illegible]

CGAGATCTAAAGAGGCACCGAAGCCCTCAGGGAGTGGCTGACAAAGCTGAGGAAGG 784

TTTGTTCCTCCAGAACAACTTGCGACAGCTCTACTCTGAATTGGATCAGGCCAAGT 844  
GCGACAAAGCAAGAGNAGGACAAAGGACGACGAGAGAAAGGAGGAGCAGCAGCAAGG 16145

GGAGCAGACGAGAGAGCGGGAAGAGCAGACGAGAAGGAGGAGAGAGCAGC 16085

AGAGACCTGAAAGTCAGCCCAAGAGGACCTACAGAGTCCCTGACCAAGGAAATCATGAGCCTGA 904  
 CGACGAGAGGAGAACCGGACCGGCAAGGAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGG 16025

AAAGAGCTACGATGCTGCAGGA 929

GAAAGGAGAAGGAAGAAGCAGACAGAGG A 10000

RESULT 18  
AC073495/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC073495 318930 bp DNA linear HTG 15-OCT-2000  
Mus musculus chromosome 11 clone RP23-345K4, WORKING DRAFT  
SEQUENCE, 64 unordered pieces.  
AC073495  
AC073495.12 GI:10801940  
HTG, HTGS, PHASE1, HTGS\_DRAFT.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 318930)  
Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,  
Deckerich, D., Thomas, S., Okwuonu, G., Carlock, C., Garner, T.,  
Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,  
Buday, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,  
Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,  
Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R.,  
Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hoques, M.,  
Kosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,  
Kovari, C., Liu, J., Liu, W., Louised, H., Lozano, R.J., Martin, R.,  
Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,  
Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B.,  
Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S.,  
Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A.,  
Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G.,  
Worley, K. and Gibbs, R.  
Direct Submission  
Unpublished  
2 (bases 1 to 318930)  
Worley, K.C.  
Direct Submission  
Submitted (19-JUN-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Oct 15, 2000 this sequence version replaced gi:10800205.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: MACH  
Center clone name: RP23-345K4  
----- Summary Statistics  
Sequencing vector: MJ3, 108821  
Chemistry: Dye-Primer Bodypy, 71% of reads  
Chemistry: Dye-terminator Big Dye, 29% of reads  
Assembly program: Phrap, version 0.990329  
Consensus quality: 282577 bases at least Q40  
Consensus quality: 299603 bases at least Q30  
Consensus quality: 309271 bases at least Q20  
Estimated insert size: 298164; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-gel estimation  
Quality coverage: 5.1x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_dirft.data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_dirft.data.html))  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 64 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 100449: contig of 100449 bp in length  
\* 100450 100549: gap of unknown length  
\* 100550 143847: contig of 43298 bp in length  
\* 143848 143947: gap of unknown length

143948 152393: contig of 8446 bp in length  
\* 152394 152499: gap of unknown length  
\* 152499 160427: contig of 7933 bp in length  
\* 160427 160526: gap of unknown length  
\* 160527 166643: contig of 6117 bp in length  
\* 166644 166744: gap of unknown length  
\* 166744 173484: contig of 6741 bp in length  
\* 173485 173585: gap of unknown length  
\* 173585 177408: contig of 3824 bp in length  
\* 177409 177509: gap of unknown length  
\* 177509 183311: contig of 5802 bp in length  
\* 183311 183410: gap of unknown length  
\* 183411 188977: contig of 5567 bp in length  
\* 188977 189078: gap of unknown length  
\* 189078 189629: contig of 7552 bp in length  
\* 189630 196729: gap of unknown length  
\* 196730 202134: contig of 5405 bp in length  
\* 202135 202235: gap of unknown length  
\* 202235 206963: contig of 4729 bp in length  
\* 206964 207063: gap of unknown length  
\* 207064 212314: contig of 5251 bp in length  
\* 212315 212414: gap of unknown length  
\* 212415 216415: contig of 4001 bp in length  
\* 216416 216515: gap of unknown length  
\* 216516 219625: contig of 3110 bp in length  
\* 219626 219725: gap of unknown length  
\* 219726 224004: contig of 4279 bp in length  
\* 224005 224104: gap of unknown length  
\* 224105 227363: contig of 3259 bp in length  
\* 227364 227463: gap of unknown length  
\* 227464 232010: contig of 4547 bp in length  
\* 232011 232110: gap of unknown length  
\* 232111 234955: contig of 2845 bp in length  
\* 234956 235055: gap of unknown length  
\* 235056 238628: contig of 3573 bp in length  
\* 238629 242373: gap of unknown length  
\* 242374 242473: gap of unknown length  
\* 242474 244574: contig of 2001 bp in length  
\* 244575 244575: gap of unknown length  
\* 244576 247650: contig of 3076 bp in length  
\* 247651 247750: gap of unknown length  
\* 247751 250706: contig of 2956 bp in length  
\* 250707 250806: gap of unknown length  
\* 250807 253646: contig of 2840 bp in length  
\* 253647 253746: gap of unknown length  
\* 253747 255897: contig of 2151 bp in length  
\* 255898 255997: gap of unknown length  
\* 255998 258504: contig of 2507 bp in length  
\* 258505 258604: gap of unknown length  
\* 258605 260676: contig of 2072 bp in length  
\* 260677 260776: gap of unknown length  
\* 260777 263190: contig of 2414 bp in length  
\* 263191 263290: gap of unknown length  
\* 263291 264984: contig of 1694 bp in length  
\* 264985 265084: gap of unknown length  
\* 265085 266722: contig of 1638 bp in length  
\* 266723 266822: gap of unknown length  
\* 266823 268472: contig of 1650 bp in length  
\* 268473 268572: gap of unknown length  
\* 268573 270376: contig of 1804 bp in length  
\* 270377 270476: gap of unknown length  
\* 270477 272058: contig of 1582 bp in length  
\* 272059 272158: gap of unknown length  
\* 272159 274342: contig of 2084 bp in length  
\* 274343 274442: gap of unknown length  
\* 274443 276862: contig of 2520 bp in length  
\* 276863 276962: gap of unknown length  
\* 276963 279114: contig of 2152 bp in length  
\* 279115 279214: gap of unknown length  
\* 279215 280875: contig of 1661 bp in length  
\* 280876 280975: gap of unknown length  
\* 280976 282744: contig of 1769 bp in length



```
* 41020 41343: gap of unknown length
* 41344 59365: contig of 18022 bp in length
* 59366 59689: gap of unknown length
* 59690 73640: contig of 13951 bp in length
* 73641 73964: gap of unknown length
* 73965 87377: contig of 13412 bp in length
* 87377 87700: gap of unknown length
* 87701 101024: contig of 13324 bp in length
* 101025 101349: gap of unknown length
* 101349 112980: contig of 11632 bp in length
* 112981 113304: gap of unknown length
* 113305 124819: contig of 11515 bp in length
* 124820 125143: gap of unknown length
* 125144 136091: contig of 10948 bp in length
* 136092 136415: gap of unknown length
* 136416 144125: contig of 7710 bp in length
* 144126 144449: gap of unknown length
* 144450 152005: contig of 7556 bp in length
* 152006 152329: gap of unknown length
* 152330 159587: contig of 7258 bp in length
* 159588 159911: gap of unknown length
* 159912 167133: contig of 7223 bp in length
* 167133 167458: gap of unknown length
* 167459 173757: contig of 6299 bp in length
* 173758 174081: gap of unknown length
* 174082 179850: contig of 5769 bp in length
* 179851 180174: gap of unknown length
* 180175 185555: contig of 5381 bp in length
* 185556 185879: gap of unknown length
* 185880 191144: contig of 5265 bp in length
* 191145 191468: gap of unknown length
* 191468 196412: contig of 4945 bp in length
* 196413 196735: gap of unknown length
* 196736 201258: contig of 4523 bp in length
* 201259 201581: gap of unknown length
* 201582 205749: contig of 4168 bp in length
* 205750 206072: gap of unknown length
* 206073 209959: contig of 3887 bp in length
* 209960 210282: gap of unknown length
* 210283 213540: contig of 3258 bp in length
* 213541 213863: gap of unknown length
* 213864 216191: contig of 2328 bp in length.
FEATURES
    source          1..216191
                    /organism="Mus musculus"
                    /db_xref="taxon:10090"
                    /clone="RP23-60B7"
BASE COUNT      60680 a 45448 c 44115 g 58724 t 7224 others
ORIGIN
Query Match      2.9%: Score 59; DB 2; Length 216191;
Best Local Similarity 43.8%: Pred. No. 8.2e-05;
Matches 257; Conservative 0; Mismatches 330; Indels 0; Gaps 0;
QY 304 GAGGAGAGAGAGATGTTGATCGAATTTCTTAAGATGACATGACATGTCAGA 363
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109137 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 109078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 364 GCCAGAGCTTTCCAGAAAGCAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109077 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 109018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 424 CGGAGATGCTGGAAGAGCAATGCTACTGCTGATCTTTCAGAGAGAGAGAGAG 483
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 109017 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 484 GCCGAGATGCTGCTCCACACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108957 GAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 544 GAGACCAACCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 108897 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108838
QY 604 CAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGGTGGAGAGATGATCCGAGCANG 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108837 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108778
QY 664 GGTGTGGAGAGTCAACCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108777 GAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108718
QY 724 TACGAGATCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108717 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108658
QY 784 GATTGCTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108657 GAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108598
QY 844 TTGAACTGAGTCAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108597 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108551
RESULT 20
AF410847/c 2065 bp DNA linear VRL 20-SEP-2001
LOCUS Ovine herpesvirus 2 latency associated antigen (orf73) gene,
DEFINITION complete cds.
ACCESSION AF410847
VERSION AF410847.1 GI:15705839
KEYWORDS Ovine herpesvirus 2.
SOURCE Ovine herpesvirus 2.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae.
REFERENCE 1 (bases 1 to 2065)
AUTHORS Coulter,L.J. and Reid,H.W.
TITLE Isolation and expression of three open reading frames (ORFs) from
ovine herpesvirus 2
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2065)
AUTHORS Coulter,L.J. and Reid,H.W.
TITLE Direct Submision
JOURNAL Submitted (15-AUG-2001) Virology, Moredun Research Institute,
Pentlands Science Park, Bush Loan, Penicuik, Midlothian EH26 0PZ,
UK
FEATURES
    source          Location/Qualifiers
                    1..2065
                    /organism="Ovine herpesvirus 2"
                    /viralon
                    /db_xref="taxon:10398"
                    complement(63..1550)
                    /gene="orf73"
                    complement(63..1550)
                    /gene="orf73"
                    /note="LANA: similar to HVS and HHV-8 Orf73"
                    /codon_start=1
                    /product="latency associated antigen"
                    /protein_id="AA05844.1"
                    /db_xref="GI:15705840"
                    /translation="MYLRSSTSDDEDDGRRRPQPKRPVTEGEGSGEGEEBPG
                    GEGEGEGVEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEG
                    VEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEG
                    GPEGBEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEG
                    GEGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEGPGEGEG
                    EEEEBEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEG
                    EEPGEGEGPGEGEGPGEGEGEGEEDPEDEPGSPSGPVRGRKRPKHOPETBRAK
                    RKKLAPIMPTLKEATYSLHNLCTSKDPVRSVRSVRALPNPAHSNIFFTGMYTFV
                    IYGNKKEATVESLFPLODAMNMPDAGANISTGTPLTSLPFPNQ"
BASE COUNT      351 a 684 c 303 g 727 t
ORIGIN
```

Query Match 2.9%; Score 58.8; DB 14; Length 2065;  
 Best Local Similarity 44.5%; Pred. No. 5e-05;  
 Matches 277; Conservative 0; Mismatches 342; Indels 3; Gaps 1;

```

OY 309 GGAGGAGATGCTTGGATCGAGAAATTTTAAAGATGAATGACAAATGTCAGAGCCCA 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1337 GGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 369 GCTTCCCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1277 CCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 429 TAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1217 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 489 GATGCTGTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1157 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 549 CAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1097 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 606 GATTGAGCTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1037 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 978
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 666 TGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 977 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 918
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 726 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 785
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 917 GTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 786 TTTGTTTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 845
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 857 CTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 846 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 905
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 797 GGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 906 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 737 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 716
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 21  
 AC079818 239480 bp DNA linear HTG 29-JAN-2002  
 LOCUS AC079818  
 DEFINITION Mus musculus clone rp23-362c8, WORKING DRAFT SEQUENCE, 10 unordered  
 pieces.  
 AC079818  
 VERSION AC079818.29 GI:17386282  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Fu, Y. and Roe, B.A.  
 Mus musculus BAC Clone rp23-362c8  
 Unpublished  
 2 (bases 1 to 239480)  
 REFERENCE  
 AUTHORS Fu, Y. and Roe, B.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-SEP-2000) Department of Chemistry And Biochemistry,  
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA  
 On Dec 6, 2001 this sequence version replaced gl:16152280.  
 COMMENT  
 ----- Genome Center

Center: Department of Chemistry And Biochemistry  
 The University of Oklahoma  
 Center code: UOKNOR

NOTE: This is a 'working draft' sequence. It currently  
 consists of 10 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

```

1 2963: contig of 2963 bp in length
* 2964: gap of 3063 bp in length
* 3064: gap of 10829 bp in length
* 10830: gap of 10929 bp in length
* 10930: gap of 23142 bp in length
* 23143: gap of 23242 bp in length
* 23243: gap of 40574 bp in length
* 40575: gap of 40674 bp in length
* 40675: gap of 63500 bp in length
* 63501: gap of 63600 bp in length
* 63601: gap of 94656 bp in length
* 94657: gap of 94756 bp in length
* 94757: gap of 121639 bp in length
* 121640: gap of 121740 bp in length
* 121740: gap of 155608 bp in length
* 155609: gap of 155708 bp in length
* 155709: gap of 197431 bp in length
* 197431: gap of 197532 bp in length
* 197532: gap of 239480 bp in length
* 239480: contig of 41949 bp in length.
Location/Qualifiers
source 1. 239480
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="rp23-362c8"

```

BASE COUNT 59470 a 60011 c 61208 g 57870 t 921 others  
 ORIGIN

Query Match 2.9%; Score 58.8; DB 2; Length 239480;  
 Best Local Similarity 45.1%; Pred. No. 9.4e-05;  
 Matches 219; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

```

OY 304 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155054 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 364 GCCAGCTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155114 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 424 CGGATACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155174 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 484 GCCGAGATGCTGTGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155234 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155293
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 544 GAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155294 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155353
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 604 CAGATTGAGCTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155354 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155413
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 664 GGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 155414 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155473
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



[illegible]







```
misc_feature      complement(16153.6296)
                   /note="match: STS: Em: L41071"
misc_feature      19774.20548
                   /note="Cpg Island"
misc_feature      /evidence=not_experimental
                   complement(120917.21304)
                   /note="match: GSS: Em: A2837024"
misc_feature      21687.22022
                   /note="match: GSS: Em: A2791707"
misc_feature      24059.24666
                   /note="match: GSS: Em: A2303512"
misc_feature      complement(126781.27451)
                   /note="match: GSS: Em: A2027075"
misc_feature      complement(126962.27413)
                   /note="match: GSS: Em: A2260334"
misc_feature      /note="match: GSS: Em: A2260334"
                   complement(130964.31484)
                   /note="match: GSS: Em: A2303774"
misc_feature      complement(131902.31984)
                   /note="match: GSS: Em: A2791347"
misc_feature      complement(139625.40432)
                   /note="match: GSS: Em: A2711476"
misc_feature      complement(139926.40442)
                   /note="match: GSS: Em: A2711491"
misc_feature      complement(144322.44838)
                   /note="match: GSS: Em: A2264680"
misc_feature      47267.47815
                   /note="match: GSS: Em: A2753028"
misc_feature      52135
                   /note="random repeat. Forced join. Gap sized to be
                   approximately 1100bp from restriction digest data and
                   spanning p1C clone."
misc_feature      62719.63094
                   /note="match: GSS: Em: A2259817"
misc_feature      complement(163268.63929)
                   /note="match: GSS: Em: A2087144"
misc_feature      complement(163297.63853)
                   /note="match: GSS: Em: A2032738"
misc_feature      64012.64531
                   /note="match: GSS: Em: A2115082"
misc_feature      64026.64425
                   /note="match: GSS: Em: A2298969"
misc_feature      complement(178311.78793)
                   /note="match: GSS: Em: A2069693"
misc_feature      complement(193998.94487)
                   /note="match: GSS: Em: A2253564"
misc_feature      94195.94814
                   /note="match: GSS: Em: A2448780"
misc_feature      95929.96536
                   /note="match: GSS: Em: A2394052"
misc_feature      97777.98336
                   /note="match: GSS: Em: A2284926"
misc_feature      complement(101592.102154)
                   /note="match: GSS: Em: A2448471"
misc_feature      complement(101908.102030)
                   /note="match: GSS: Em: A2448470"
misc_feature      complement(103287.103341)
                   /note="match: GSS: Em: A2394388"
misc_feature      106254.107006
                   /note="Cpg Island"
                   /evidence=not_experimental
misc_feature      complement(114284.114736)
                   /note="match: GSS: Em: A2041515"
misc_feature      match: STS: Em: HS28615T"
                   complement(116196.116371)
                   /note="match: GSS: Em: A2056615"
misc_feature      complement(127255.127857)
                   /note="match: GSS: Em: A2268926"
misc_feature      127968.128455
                   /note="match: GSS: Em: A2237605"
misc_feature      complement(128635.128910)
                   /note="match: GSS: Em: A2788648"
misc_feature      143168.143545
                   /note="match: GSS: Em: A2246254"
```

BASE COUNT	ORIGIN
misc_feature	146147..146196 /note="Sequence from uni-directional primer reads and dGTP big dye terminator reads only."
misc_feature	complement(151206)..151860 /note="match: GSS: Em:A0930179"
misc_feature	complement(152098)..152415 /note="match: STS: Em:G07175"
misc_feature	complement(152794)..153079 /note="match: GSS: Em:A470074"
misc_feature	152905..153117 /note="match: GSS: Em:A2797680"
misc_feature	159586..160222 /note="match: GSS: Em:A284811"
misc_feature	165617..167046 /note="Cpg island"
misc_feature	/evidence=not_experimental complement(167307)..167566 /note="match: GSS: Em:A2784357"
misc_feature	187090..187339 /note="match: GSS: Em:A2483007"
misc_feature	189492..189624 /note="match: GSS: Em:A2846852"
BASE COUNT	56357 a 42439 c 41402 g 52477 t

[illegible]



TITLE  
JOURNAL  
COMMENT

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karats,A.,  
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoccky,J.,  
Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thoman,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 10, 2001 this sequence version replaced gi:8139804.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

## ----- Project Information

Center clone name: L9029

Sequencing vector: M13

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap; version 0.960731

Consensus quality: 150458 bases at least Q40

Consensus quality: 165983 bases at least Q20

Insert size: 153000; agarose-fp

Insert size: 169819; sum-of-contigs

Quality coverage: 4.7 in Q20 bases; agarose-fp

Quality coverage: 4.2 in Q20 b.

NOTE: This is a 'working draft' sequence. It currently

consists of 43 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 13487: contig of 13487 bp in length

13488 13587: gap of 100 bp

13588 14626: contig of 1039 bp in length

14627 14726: gap of 100 bp

14727 15793: contig of 1067 bp in length

15794 15893: gap of 100 bp

15894 17057: contig of 1164 bp in length

17058 17157: gap of 100 bp

17158 18291: contig of 1134 bp in length

18292 18391: gap of 100 bp

18392 20046: contig of 1655 bp in length

20047 20146: gap of 100 bp

20147 21386: contig of 1240 bp in length

21387 21486: gap of 100 bp

21487 22760: contig of 1274 bp in length

22761 22860: gap of 100 bp

22861 23953: contig of 1093 bp in length

23954 24053: gap of 100 bp

24054 25081: contig of 1028 bp in length

25082 25181: gap of 100 bp

25182 26357: contig of 1176 bp in length

26358 26457: gap of 100 bp

26458 27574: contig of 1117 bp in length

## FEATURES

## source

27575 27674: gap of 100 bp  
27675 28756: contig of 1082 bp in length  
28757 28856: gap of 100 bp  
28857 30554: contig of 1699 bp in length  
30555 30654: gap of 100 bp  
30655 32124: contig of 1470 bp in length  
32125 32224: gap of 100 bp  
32225 33634: contig of 1410 bp in length  
33635 33734: gap of 100 bp  
33735 35520: contig of 1786 bp in length  
35521 35620: gap of 100 bp  
35621 36833: contig of 1213 bp in length  
36834 36933: gap of 100 bp  
36934 38632: contig of 1699 bp in length  
38633 38732: gap of 100 bp  
38733 40693: contig of 1961 bp in length  
40694 40793: gap of 100 bp  
40794 41899: contig of 1106 bp in length  
41900 41999: gap of 100 bp  
42000 43910: contig of 1911 bp in length  
43911 44010: gap of 100 bp  
44011 45470: contig of 1460 bp in length  
45471 45570: gap of 100 bp  
45571 46993: contig of 1423 bp in length  
46994 47093: gap of 100 bp  
47094 48770: contig of 1677 bp in length  
48771 48870: gap of 100 bp  
48871 51099: contig of 2229 bp in length  
51100 51199: gap of 100 bp  
51200 53719: contig of 2520 bp in length  
53720 53819: gap of 100 bp  
53820 56012: contig of 2193 bp in length  
56013 56112: gap of 100 bp  
56113 57226: contig of 1114 bp in length  
57227 57326: gap of 100 bp  
57327 77403: contig of 20077 bp in length  
77404 77503: gap of 100 bp  
77504 79994: contig of 2491 bp in length  
79995 80094: gap of 100 bp  
80095 82923: contig of 2829 bp in length  
82924 83023: gap of 100 bp  
83024 86207: contig of 3184 bp in length  
86208 86307: gap of 100 bp  
86308 90457: contig of 4150 bp in length  
90458 90557: gap of 100 bp  
90558 96908: contig of 6351 bp in length  
96909 97008: gap of 100 bp  
97009 102861: contig of 5853 bp in length  
102862 102961: gap of 100 bp  
102962 107881: contig of 4920 bp in length  
107882 107981: gap of 100 bp  
107982 113343: contig of 5562 bp in length  
113344 113643: gap of 100 bp  
113644 118163: contig of 4520 bp in length  
118164 118263: gap of 100 bp  
118264 135271: contig of 17008 bp in length  
135272 135371: gap of 100 bp  
135372 151730: contig of 16359 bp in length  
151731 151830: gap of 100 bp  
151831 171146: contig of 19316 bp in length  
171147 171246: gap of 100 bp  
171247 174019: contig of 2773 bp in length.

## Location/Qualifiers

1. 174019

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-228G18"

/clone\_id="RP11 Human Male BAC"

1. 13487

/note="assembly\_fragment"

clone\_end:SP6

## misc\_feature



```
misc_feature      13588.14626      /note="assembly_fragment"  
misc_feature      14727.15793      /note="assembly_fragment"  
misc_feature      15894.17057      /note="assembly_fragment"  
misc_feature      17158.18291      /note="assembly_fragment"  
misc_feature      18392.20046      /note="assembly_fragment"  
misc_feature      20147.21386      /note="assembly_fragment"  
misc_feature      21487.22760      /note="assembly_fragment"  
misc_feature      22861.23953      /note="assembly_fragment"  
misc_feature      24054.25081      /note="assembly_fragment"  
misc_feature      25182.26357      /note="assembly_fragment"  
misc_feature      26458.27574      /note="assembly_fragment"  
misc_feature      27675.28756      /note="assembly_fragment"  
misc_feature      28857.30554      /note="assembly_fragment"  
misc_feature      30655.32124      /note="assembly_fragment"  
misc_feature      32225.33634      /note="assembly_fragment"  
misc_feature      33735.35520      /note="assembly_fragment"
```

Query Match	2.88;	Score 56.2;	DB 2;	Length 174019;
-------------	-------	-------------	-------	----------------

Matches 250; Conservative 0; Mismatches 303; Indels 2; Gaps 1.

Oy	376	CAAAAAGCAGAGAAACGAGACAGCCAGGCTATCATCTGACACATCTCGGGATACGGTG	435
Db	66805	CAGAAAGCAGAGAGAAATATGTGTGGAGCAGAGAAAGAAAGAACTGGCAGCAGCAAGGCTG	66864
Oy	436	GAAAGACCGAATGCTACTGTGGTATCTCTGCAGCAGGCGCTTGGGCAAGGCCAGATGCTG	495
Db	66855	CCGGACCAAGAAAGAAAGACTGTGGGAACAGCAGGAAGATGCAGAGACAGGAGAGAAAGATA	66924
Oy	496	TGCTCCACACTAAAAAGCAGATGTAAGTACTTAGAGCAGCAGCAGATGATGAGCCAAACAA	555
Db	66925	TGGAGCAGAGAGAGAAAGATACGGGACCAAGAGGAGATTTGGGGCAGAGAAAGATG	66984
Oy	556	GCACAAGAGAGCGCGGGCGGCTCAGAGACCAAGATGAAAGCCATGGAGAGATTGAGCTT	615
Db	66985	TGGCGGCGAGAGAAATGTCGG--GAGCAGGAAGATGTGGACAGAGAGAGGCTGCGAG	67042
Oy	616	CTACTCCAGAGCCAGCTCCCTAGAGTGGAGAGAGATGATCCGAGACATGGGCTGTGGACAG	675
Db	67043	GAGCAGAGAAACCAATGTGGGAGCAGGAGGAAAGATTCGGGATTCAGGAGCAGAAAGATG	67102
Oy	676	TCACGGGTGGAACACCTGGCTGTGTACTGTGTCTCTCAAAAGAGATGACGAGATCTA	735
Db	67103	TGGGACACAGAGAGAGATGTGTGGGAGGACGAGACAGAGAGGCTCGGGAGAAAGAGGAGAA	67162
Oy	736	AAAGAGCCACGAGAGGCTCAGGAGAGGTGGCTGACAAAGCTGAGGAGAGATTGTTTTCC	795
Db	67163	ATGCGGGAGCAAAAGATGTGGCAGCAGGTGGGAAAGATTCGGGAGGAGGAAGAAAGAG	67222
Oy	796	TCCAAAGACATGTGCAGACAGTCTACTCTGTAATTGGATCAGGCCCAAGTTAGACAGAG	855
Db	67223	CAGAGCAGAGAGAAAGACATGAGGAGCAGGAGAGAGAAAGATCCGAGAGGAGAGAGATG	67282
Oy	856	TCAGCCCAAGAAAGACTTACAGATGCTGTACAAGAAATCATGAGCCTGTAAAAAGAGCTTA	915
Db	67283	CGGGAGCGGGAACAAAGATGTGGGAGGAGGAGAGATATATCCGGAGCCAGGAGAGGAGAG	67342

```

QY      916  ACGATGCTGCAGGAA  930
          | | | | | | | |
Db 67343  ATGCAGGAGCAGGAA  67357

```

RESULT	29				
AC073446/c					
LOCUS	AC073446	185497 bp	DNA	linear	HTG 11-OCT-2001
DEFINITION	Homo sapiens chromosome 15 clone RP11-757E13 map 15, *** SEQUENCING				
IN PROGRESS	***, 5 unordered pieces.				

```

VERSION AC073446.7 GI:16041356
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN
SOURCE human.
ORGANISM Homo sapiens

```

REFERENCE	1 (bases 1 to 185497)
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE	Homo sapiens chromosome 15, clone RP11-757E13
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 185497)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Rodgers, S., Baidinov, J., Bana, N., Bastien, Y., Beaud, E., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campoliano, A., Castle, A., Cheopel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Deavellano, K., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzgerald, W., Gage, D., Galaun, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Hesford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kamp, L., Karatas, A., Klein, J., Lacroque, K., Iamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuck, A., McKernan, K., McPheters, N., Meldrum, J., Menues, L., Mihova, T., Miranda, C., Menga, Y., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivier, T. M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Rhomon, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

**JOURNAL** Submitted (18-JUN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**COMMENT** On Oct 11, 2001 this sequence version replaced gi:15718563.

and repeats were identified using RepeatMasker.  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www.seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L10251  
 Center clone name: 757\_E\_13

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

*	1	33974:	contig of 33974 bp in length
*	33975	34074:	gap of 100 bp
*	34075	115332:	contig of 81258 bp in length
*	115333	115432:	gap of 100 bp
*	115433	117817:	contig of 2385 bp in length

```

FEATURES
source
* 117818 117917: gap of 100 bp
* 117918 165496: contig of 47579 bp in length
* 165497 165596: gap of 100 bp
* 165597 185497: contig of 19901 bp in length.
location/qualifiers
1. 165497

```

	clone_11b-"RPCI-11	Human Male BAC"
BASE COUNT	51232 a 39784 c 40726 g 53302 t	453 others
ORIGIN		

Query Match	2.8%	Score	56.2	DB 2	Length	185497
Best Local Similarity	45.0%	Pred. No.	0.00049			
Matches 250; Conservative	0	Mismatches	303	Indels	2	Gaps 1

QY 376 CAGAAAGCAGGAGAAACGAGACAGCCAGCTCATCTGCACACTCTGGCGGATPACGCTG 435

Db 107996 CAGAAAGCAGGAGGAAATATGTGGAGCACAGAAAGAGATGTGCACGACGACGAGGCTG 1079337

QY 436 GAAGAACGCATGCTACTGTGTATCTCTCAGCAGGCGCTTGGGCAAGCGCCGATGCTG 495

Db 107936 CCGGAACGAAAGAGAAAGTGTGTGGAAACAGAGAAAGATGACAGGACGAGGAGAAAGATA 1078777

QY 496 TGCTCCACACTGAAAAAGCAGATGTAAGTACTTAgAGCAGCAGCAGATGTAGACCAACAA 555

Db 107876 TGGGAGCAGGAGGAGAAAGATACGGGACCAGAGAGAGATGTGGGGGCGAGGAAAGATG 1078171

QY 556 GCACAGAGGAGGCGCGGCGCTAGGAGCAAGTGAAGCAATGGAGCAGATTTGAGCTT 615

Db 107816 TGGCGGCAAGGGAAGATGCGG--GAGCAGGAAGTGTGGGACACAGGAGGAGAGCTGCAG 107759

QY 616 CTACTCCAGAACCCAGTCCCTGAGAGTGAGAGAGATGATCCGAGCATAGGGTGTGGGACAG 675

Db 107758 GAGCAGGAGGAACGAGATGTGGGAGACAGAGAGAGAAAGATGGGGATTCAGAGCAGAAAGATG 1076939

QY 676 TCAGCGGTGCAACGCTGGCTGTGTACTGTGTCTCTCAAGAAAGATACGAAATCTA 735

Db 107698 TGGGACACAGGAGGAGAGATGTGGGACACAGAGCAGAGGCTGCGGGAGAGAGAGAGAGA 1076339

QY 736 AAAAGAGCAGGAGGCGCTCAGGAGAGTGGCTGACAACTGAGAGAAAGATTTGTTTCC 795

Db 107638 ATGCGGGAGCGAAGAAAGATGTGGCACAGGTGAGAAAGATGCGGGAGGAGAAAGAACG 107579

QY 796 TCCAGAAAGCAGTTCGACAGACTGTACTTGAATTGGATCAGGCCAAGTTAGAACTGAAG 855

Db 107578 CAGGAGCAGGAGAAAGAAAGACTGTGGACACAGAGAGAGAAAGATGCCAGAGGAGAGAGATG 1075139

QY 856 TCACCCCGAAGAGACTTACAGAGTGTGACAAAGGAAATCATGAGCCTGAAAAAGAGCTA 915

Db 107518 CGGAGCGGGAGAGAAAGATGTCGGAGAGAGAGAGATGATCGGGAGCAGAGAGAGAG 107459

QY 916 ACGATGCTGCAGGAA 930

Db 107458 ATGCAAGGAGCAGGAA 107444

RESULT	30
AL606962	
LOCUS	223426 bp DNA linear HTG 30-JAN-2002
DEFINITION	Mus musculus chromosome 4 clone RP23-29H22, *** SEQUENCING IN PROGRESS ***, In unordered pieces.
ACCESSION	AL606962
VERSION	AL606962.14 GI:18476783
KEYWORDS	HTGS; HTGS_PHASE1; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	house mouse
ORGANISM	Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (sites)  
Babbage, A.  
Direct Submission  
Submitted (29-JAN-2002) Wellcome Trust Sanger Institute, Hinxton

COMMENT On Feb 1, 2002 this sequence version replaced gi:18250530.

- \* This record will be updated with the finished sequence
- \* as soon as it is available and the accession number will
- \* be preserved.

FEATURES	
source	location/Qualifiers
	1. .223426
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/chromosome="4"
	/clone="RP23-29H22"
	/clone_lib="RPC1-23"
misc_feature	1. .52434
	/note="assembly_fragment:02463.0"
misc_feature	52535. .195602
	/note="assembly_fragment:02463.1"
misc_feature	195703. .223426
	/note="assembly_fragment:05002"
BASE COUNT	63793 a 51782 c 50189 g 57462 t 200 others
ORIGIN	

Query Match	2.8%	Score	56.2	DB 2	Length	223426
Best Local Similarly	45.7%	Pred. No.	0.0005			
Matches 196	Conservative	0	Mismatches	233	Indels	0
					Gaps	0

[illegible]

Db	34963	GGAGAAAGGAGAAGGAAGACAGACAGACAAGAAAGAAAGAAAGGAAGAACAGACAGACA	35022
Qy	870	CITACAGACTGCTGCACAGCAATATCATGACCGCTGAAAAAAGACTTACAGACTCTCAGCA	929
Db	35023	GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	35082
Qy	930	AACCTTGAA	938
Db	35083	GAGAAGAA	35091
<hr/>			
RESULT	31		
AL645746			
DEFINITION	Mus musculus chromosome 13 clone RP23-153B6, *** SEQUENCING IN		
ACCESSION	AL645746		
VERSION	AL645746.2		
KEYWORDS	HTG: HTGS, PHASE1.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1 (sites)		
TITLE	Sims, S.		
COMMENT	Direct Submission Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Nov 28, 2001 this sequence version replaced gi:17017892. ----- Genome Center ----- Center: UK Medical Research Council Center code: UK-MRC Web site: http://micseq.har.mrc.ac.uk Contact: mouseq@har.mrc.ac.uk ----- Project Information ----- Center project name: Bm153B6 ----- Summary Statistics ----- Assembly program: XGAP4; version 4.5 Sequencing vector: plasmid; 108752; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Consensus quality: 279809 bases at least Q40 Consensus quality: 280516 bases at least Q20 Consensus quality: 281062 bases at least Q20 Insert size: 281711; sum-of-contigs Insert size: 190968; 4.7% error; agarose-fp Quality coverage: 12.39x in Q20 bases; sum-of-contigs Quality coverage: 18.52x in Q20 bases; agarose-fp		
<hr/>			
* NOTE: This is a 'working draft' sequence. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.			
FEATURES			
source	location/Qualifiers		
	1..282611	/organism="Mus musculus"	
	/db_xref="taxon:10090"		
	/chromosome="13"		
	/clone="RP23-153B6"		
	/clone_lib="RPCR-23"		
	1..30613		
misc-feature	/note="assembly_fragment:06533 fragment_chain:1"		
misc-feature	30714..46652		
misc-feature	/note="assembly_fragment:02685 fragment_chain:1"		
misc-feature	46753..51050		
misc-feature	/note="assembly_fragment:08628 fragment_chain:1"		
misc-feature	51151..57360		
misc-feature	/note="assembly_fragment:06992 fragment_chain:1"		
misc-feature	57461..61884		

[illegible]

FEATURES	source	Location/Qualifiers
JOURNAL	herpesvirus encoding a protein with similarities to latent membrane proteins 1 and 2A of Epstein-Barr virus	
MEDLINE	J. Virol. 73 (8), 6953-6963 (1999)	
REFERENCE	2. (bases 1 to 28559)	
AUTHORS	Glenn, M.A., Rainbow, L., Aurade, F., Davison, A.J. and Scholz, T.F.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-AR-1999) MRC Virology Unit, Church Street, Glasgow G11 5JR, UK	
gene		1. .28559
CDS		/organism="Human herpesvirus 8"
		/isolate="GK18"
		/db_xref="taxon:37296"
		/note="right end of the unique region plus 8 complete copies and 1 partial copy of the terminal repeat (TR)"
		<1. .2105
		/gene="ORF 68"
		<1. .1102
		/gene="ORF 68"
		/function="involved in DNA packaging"
		/note="similar to HSV-1 UL31, HCMV UL52 and EBV BFLF1"
		/codon_start=2
		/evidence=not_experimental
		/protein_id="AAd6496.1"
		/db_xref="GI:5669899"
		/translation="IMHLHPEEATYGLPGRGLMGIDLIHFVQKCFKTAAKI LGISNLOEIKSEFRCMLTGITICNPFCKSPMDPEKATGPPCOIDTTPAPAS GIEPLARATFCGASRPTRPSILRALIDIMWTSSELDDEPRRLIASMSBELKSVASH DPEFSPPLADTSQGPCILMHPTGLRKNGTASVCLDECLAHPEAPKALQTLQCEVH MGHEENNVKLVDRIAEVLDNPFAPYSDPLRELILGCTPOEIHKLHCPDLICALNA KVSSEVFLRLPREOEYKKLRASAAAGOLDANTLPDCEVQGLVLFKGIQNNRVGK TTSIDILRELTAQKKRRRLDLAHSQTSHLXA"
		1135. .2105
gene		/gene="ORF 69"
CDS		1135. .2043
		/gene="ORF 69"
		/note="similar to HSV-1 UL31, HCMV UL53 and EBV BFLF2"
		/codon_start=1
		/evidence=not_experimental
		/protein_id="AAd6497.1"
		/db_xref="GI:5669890"
		/translation="MPKSVSSHLSLATSTGRSGPRDIRCLSRLSRSPVPGARSASVS SKHNGLRKPIISDKVFPSILSHRELGVDFIREKEPTICTSYMLPLDLSIVPGRC VLSPPGHSGMFQCALCPSTENPTVAOGRPOTMGDLKKNELCSVALAEVHHNA DKVIOKHTFYLISLSHMDVYROSFIQGLYANLYLKTGHDPILFTTNNGLTJWC ILFTRRLHIGETRLRLMDNLPNYKISADCCROSYVVKFVPTHPDASTIAVQVHTTC EAAVALCDTDEMDDIDKGTALVNAL"
		2100. .2105
polyA_signal		/gene="ORF 69"
		/note="possible signal for ORFs 68 and 69"
		/evidence=not_experimental
		complement(2126)
polyA_site		/gene="ORF K12"
		/evidence=experimental
		complement(2126. .2824)
mrna		/gene="ORF K12"
		/product="T0.7"
		/evidence=experimental
		complement(2126. .2824)
gene		/gene="ORF K12"
		complement(2139. .2144)
polyA_signal		/gene="ORF K12"
		/evidence=experimental
		complement(2616. .2798)
CDS		/gene="ORF K12"
		/note="transforming gene"
		/codon_start=1
		/evidence=experimental
		/product="Kapsin"
		/protein_id="AAd6499.1"
		/db_xref="GI:5669892"
		/translation="MDRGITVFVAVHVDVLLNGMRMRGAIPPLVCLLISVPPSGS

[illegible]

[illegible][illegible]

Query Match	2.8%	Score 55.4	DB 14	Length 3127
Best Local Similarity	45.6%	Pred. No. 0.00047		
Matches 194	Conservative 0	Mismatches 231	Indels 0	Gaps 0
QY	505	CTGAAAGACGATGMACTACTTTAGACGACGACGAGATGAGACCAACACACAAAG	564	
Db	1738	CAGCAGGATGAGACACACAGATGAGCAGCAGCAGCAGGATGAACAGGACGACGAG	1797	
QY	565	GAGCGGGCCGGCTTCAGACGAAGATGAAGACCATGAGACAGATTGACTTCTACTCCAG	624	
Db	1798	GAGCGAGGACGACGAGGAGACGAGGAGGAGGATTAGAGAGACGACGACGAGTTAGAG	1857	
QY	625	AGCCAGCTCCCTGAGAGTGGAGAGATGATCCGACATATGGGTGTGGACATGACCGGTG	684	
Db	1858	GAGCAGGAGACGAGGATTTAGAGAGCAGCAGGAGAGTGTAGAGACAGGACGAGAGTTA	1917	
QY	685	GAAACAGTGGCTGTGTACTGTGTGTCTCTCAAGAAAGAGTACGAGATCTAAAGAGCA	744	
Db	1918	GAGGAGCAGGAGACAGAGATTAGAGGACGACGAGGAGTTAGAGACGACGAGCAGAG	1977	
QY	745	CGGAAGGCGCTCAGGAGGAGTGGCTGCACAAAGCTGAGGAGAGATTGTTTTCCTCCAGAGC	804	
Db	1978	TTAGAGGAGCAGGACGACAGAGATTAGATGAGCAGGAGCAGGACTTAGAGAGCAGAGCAG	2037	
QY	805	AAGTTCGACAGACAGTCTACTCTGTAATTTGGATCAGGCCAAGTTAGAACTGAAGTCAGCCCG	864	
Db	2038	GAGTTAGAGGAGCAGCAGGACGAGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGAGCAGAG	2097	
QY	865	AAGGACTTCACAGAGTGTGCACAGGAATCATGACCTGTGAAAAAGAGTAAACGATGCTG	924	
Db	2098	CAGGAGTTAGAGGACGACGAGGACGAGGAGTTAGAGGAGCAGGACGAGTTAGAGAGCAG	2157	
QY	925	CAGCA 929		
Db	2158	GAGCA 2162		
RESULT 34				
AC098712	207683 bp	DNA	Linear	HTG 14-NOV-2001
LOCUS				
DEFINITION	Mus musculus chromosome UNK clone RP23-2A21, WORKING DRAFT			
ACCESSION	AC098712.2	GI:16924164		
VERSION	AC098712.2	GI:16924164		
KEYWORDS	HTG, PHASE1, HTGS_DRAFT, HTGS_ACTIVEFIN,			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (bases 1 to 207683)			
JOURNAL	McPherson, J.D. and Waterston, R.H.			
REFERENCE	The sequence of Mus musculus clone			
AUTHORS	Unpublished			
TITLE	2 (bases 1 to 207683)			
JOURNAL	McPherson, J.D. and Waterston, R.H.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (31-Oct-2001) Genome Sequencing Center, 4444 Forest Park			
TITLE	Parkway, St. Louis, MO 63108, USA			
JOURNAL	On Nov 14, 2001 this sequence version replaced gi:16554409.			
COMMENT				
	Genome Center			
	Center: Washington University Genome Sequencing Center			
	Center code: WUGSC			
	Web site: http://genome.wustl.edu/gsc/index.shtml			
	Contact: submissions@wustl.edu			
	Project Information			
	Center project name: M.BA0002A21			
	Summary Statistics			
	Sequencing vector: M13: &			
	Sequencing vector: plasmid: &			

				Chemistry: Dye-primer ET: % of reads
				Chemistry: Dye-terminator Big Dye: % of reads
				Assembly program: Phrap; Version 0.990319
				Consensus quality: 205930 bases at least Q40
				Consensus quality: 206594 bases at least Q30
				Consensus quality: 207083 bases at least Q20
				Insert size: -; agarose-fp
				Insert size: 21954; sum-of-contigs
				Quality coverage: 25.41 in Q20 bases; agarose-fp
				Quality coverage: 24.42 in Q20 bases; sum-of-contigs
				-----
				* NOTE: This is a 'working draft' sequence. It currently
				* consists of 5 contigs. The true order of the pieces
				* is not known and their order in this sequence record is
				* arbitrary. Gaps between the contigs are represented as
				* runs of N, but the exact sizes of the gaps are unknown.
				* This record will be updated with the finished sequence
				* as soon as it is available and the accession number will
				* be preserved.
				1 1240: contig of 1240 bp in length
				1241 1340: gap of unknown length
				1341 14229: contig of 12889 bp in length
				14230 14329: gap of unknown length
				14330 39801: contig of 25472 bp in length
				39802 39901: gap of unknown length
				39902 71680: contig of 31779 bp in length
				71681 71780: gap of unknown length
				71781 207683: contig of 135903 bp in length.
				Location/Qualifiers
				1..207683
				/organism="Mus musculus"
				/db_xref="taxon:10090"
				/chromosome="GNK"
				/clone="RP23-2A21"
				-----
				misc_feature
				1..1240
				/note="assembly_name:Contig14"
				1341..14229
				/note="assembly_name:Contig37"
				14330..39801
				/note="assembly_name:Contig38"
				39902..71680
				/note="assembly_name:Contig39"
				71781..207683
				/note="assembly_name:Contig40"
				BASE COUNT 62336 a 44191 c 44243 g 56460 t 453 others
				ORIGIN
				-----
				Query Match 2.8% Score 55.4; DB 2; Length 207683;
				Best Local Similarity 44.7%; Pred. No. 0.00083;
				Matches 215; Conservative 0; Mismatches 266; Indels 0; Gaps 0;
				-----
OY	305	AGGAGGAGGAGCAATCTTGGATCGAGATTCTTAAAGATGAACCTGGACATGTCAGAG	364	
DB	186130	AGGAGCAGAGAGAGCAGGAGAGGAGGAGAGGAGAGAGAGAGAGAGAGAGCAGGAGAGC	186189	
OY	365	CCCAGCTTCCCAAGAAAGCAAGAGAAACGAGCAGCAGGCTCATCTGCACACTCTGC	424	
DB	186130	AGGAGGAGACAGC	186249	
OY	425	GGGATACGCTGGAAAGCAAGCATGCTACTGTGTATCTCTGACAGCAGGCTTGGCCAAG	484	
DB	186250	AGGAG	186309	
OY	485	CCGAGATGCTGTGCTCCACACTGAAAGAGAGATGAAGTACTTTAAGCAGCAGCAGATG	544	
DB	186310	AGGAGAGACAG	186369	
OY	545	AGACCAACCAAGCAGACAG	604	
DB	186370	AGCAG	186429	
OY	605	AGATTGACCTTCTACTCCAGAGCCAGCTCCCTGAGCTGAGAGAGATGATCCGAGACATGG	664	

[illegible]

```

/chromosome="19"
/clone="LINLR-277D11"
Lawrence Livermore human cosmid library LINLR"
BASE COUNT      7758 a 11001 c 10770 g 6119 t      200 others
ORIGIN

Query Match      2.7%: Score 55; DB 2; Length 35848;
Best Local Similarity 49.8%: Pred. No. 0.0085;
Matches 139; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY  507  GAAAAAGCAGATGATGAGTCTTAGACGACGACGAGGATGAGACCAAAACAGCAGAGAGA 566
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  19523 GGAGGAAACGAGGAGAGAGAGAGAGAGAGAAACAGGAGAGAGGAGGAGAGCAAGAGAGA 19582
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  567  GCGCGGCGCGCTCAGACAGCAGATGAAGCCATGAGCAGATTTGAGCTTCTACTCCAGAG 626
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  19583 GGAGGAGGAGGAAACAGCAGCAGAGGAGAGAGAGCCACAGGAGGAGGAGGAGGAAGCAGAGAGA 19642
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  627  CCAGCTCCCTGAGTGGAGGAGAGATGATCCGACACATGGGTGTGGACAAGTCAGCGGTGGA 686
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  19643 GGAGGAGGAAACAGGAGAGAGAGAGAGAAACAGGAGGAGGAGGAGGAACAGAGAGAGAGA 19702
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  687  ACAGCTGGCTGTGTACTGTGTGTCTCAAGAAAGATPACGAATCTTAAAGAGGCCACG 746
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  19703 GGAAACAGGAGGAGGAGAGAGAGAGAGCAGCAGGAGGAGGAGGAGGAGGAGCAGAGAGAGA 19762
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  747  GAAAGCCTCAGGAGGAGGTGGCTGACAAACCTGAGGAAAGGA 785
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  19763 GGAGGAGGAAACAGGAGAGAGAGAGAGAGAAACAGGAGAGAGA 19801
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT	36
AC091454.2/c	
MPCOMMENT	
Sequence split into 4 fragments	
Fragment Name	Begin End Locus AC091454 Accession AC091454
AC091454_0	1 110000
AC091454_1	100001 210000
AC091454_2	200001 310000
AC091454_3	300001 394326
Continuation (3 of 4) of AC091454	from base 200001 (AC091454 Mus musculus chromosome)
Query Match	2.7% Score 54; DB 2; Length 110000;
Best Local Similarity	44.1%; Pred No. 0.0019;
Matches 225; Conservative	0; Mismatches 285; Indels 0; Gaps 0;
OY	305 AGAGGAGGACAATGTCTTGGATCGAATTCTTTAAAGATGAACATGTGCACATGTGCAG 364 
Bd	69201 AGGACGAGAGCAGCAGGAGGAGGACGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 69142 
OY	365 CCCACTTTCCCAAGAAGCAAGGAGAAACGAGACGCCAGTCATCATCTGCACACTCTGC 424 
Bd	69141 AGGACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 69082 
OY	425 GGGAATACGCTTGAAAGAACCATATGCTACTGTGGTATCTCTCGACAGCGCCTTGCGCAAG 484 
Bd	69081 AGCAGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 69022 
OY	485 CCGAGATGCTGTGCTCCACACTGTAAGAAAAGCAGATGAAGTACTTAGAGCAGCAGGAGTATG 544 
Bd	69021 AGG 68962 
OY	545 AGACCAACCAAGCACAAAGGAGGAGCGCGCGGCTCAAGAGCCAAGATGAAGACCATGAGAC 604 
Bd	68961 AGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 68902 
OY	605 AGATTGAGCTTCTACTCCAGAGCCAGACTCCTTGAGGTGAGGAGATGATCCGAGCATGAG 664 
Bd	68901 AAGAGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 68842 
OY	665 GTGTGGAGCAGTCAACGCGGTGGAACAGCTGGCTGTGTACTGTGTGTCTTCGAAGAAAGAGT 724 



Db 68841 AGCAGAGCGAGCAGAGGAGCATCAGAGAGCAGCAGCAGCAGAGCAGG 68782  
QY 725 ACGAGATCTAATAAGAGCGCAGGAGGCCCTCAGGAGGTGCTGCACACACTGAGGAGG 784  
Db 68781 AGGAGGAGCAGGAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGCAGCACC 68722  
QY 785 ATTTGTTCCTCCAGACGCAAGTTGCAGA 814  
Db 68721 ATGAGGAGCGGAGACAGAGACACAGAGA 68692

RESULT 37  
AC007461 180385 bp DNA linear PRI 04-JUN-1999  
LOCUS Homo sapiens chromosome 17, clone 84\_E\_24, complete sequence.  
DEFINITION AC007461  
AC007461 GI:4996920  
VERSION HTG.  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 180385)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens chromosome 17, clone 84\_E\_24  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 180385)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,  
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,  
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,  
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grunt,G.,  
Hagos,B., Heath,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,  
Karatas,A., Lehocsky,J., Lieu,C., Locke,K., Macdonald,P.,  
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,  
Meldrum,J., Molla,M., Morris,M., Morrow,J., Mychaleckyj,J.,  
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,  
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,  
Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,  
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.  
JOURNAL Direct Submission  
TITLE Submitted (04-MAY-1999) Whitehead Institute/MIT Center for Genome  
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA  
REFERENCE 3 (bases 1 to 180385)  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,  
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,  
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,  
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grunt,G.,  
Hagos,B., Heath,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,  
Karatas,A., Lehocsky,J., Lieu,C., Locke,K., Macdonald,P.,  
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,  
Meldrum,J., Molla,M., Morris,M., Morrow,J., Mychaleckyj,J.,  
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,  
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,  
Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,  
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.  
JOURNAL Direct Submission  
TITLE Submitted (04-JUN-1999) Whitehead Institute/MIT Center for Genome  
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Jun 4, 1999 this sequence version replaced gi:496400.  
All repeats were identified using RepeatMasker: Smit, A.F.A. &  
Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Only 180385 base pairs from the middle of this clone are being  
submitted. The remainder overlaps either accession number AC005152  
(WICGR project l335) or accession number AC005243 (WICGR project

FEATURES L343).  
source  
Location/Qualifiers  
1. 180385  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="84\_E\_24"  
/clone\_lib="Alan Buckler -- per comm"  
446..753  
/rpt\_family="AluSc"  
repeat\_region  
2375..2419  
/rpt\_family="(TTTTC)n"  
repeat\_region  
complement(2985..3103)  
/rpt\_family="L2"  
repeat\_region  
complement(3146..3352)  
/rpt\_family="L2"  
repeat\_region  
complement(3531..3639)  
/rpt\_family="L2"  
repeat\_region  
complement(3798..4394)  
/rpt\_family="L2"  
5531..5631  
/rpt\_family="MER81"  
repeat\_region  
5648..5674  
/rpt\_family="(CAAAA)n"  
repeat\_region  
5812..6020  
/rpt\_family="LIMC4"  
6180..6224  
/rpt\_family="AT-rich"  
6303..6611  
/rpt\_family="AluSp"  
7378..7618  
/rpt\_family="MIR"  
9593..9623  
/rpt\_family="(GGCA)n"  
repeat\_region  
complement(10666..10923)  
/rpt\_family="AluSq"  
10958..11088  
/rpt\_family="MIR"  
11752..11914  
/rpt\_family="L2"  
repeat\_region  
complement(11915..12219)  
/rpt\_family="AluSx"  
12220..12347  
/rpt\_family="L2"  
repeat\_region  
13110..13182  
/rpt\_family="GA-rich"  
13526..13564  
/rpt\_family="(TCC)n"  
13575..13660  
/rpt\_family="(TA)n"  
complement(14380..14658)  
/rpt\_family="AluJo"  
14742..14830  
/rpt\_family="(TA)n"  
14860..14967  
/rpt\_family="(TA)n"  
complement(14977..15122)  
/rpt\_family="L2"  
repeat\_region  
complement(15292..15385)  
/rpt\_family="L2"  
16672..16781  
/rpt\_family="L2"  
repeat\_region  
16802..16846  
/rpt\_family="(TGAA)n"  
17225..17245  
/rpt\_family="(A)n"  
18411..18720  
/rpt\_family="AluJb"  
complement(20715..20788)  
/rpt\_family="MIR1J2"  
complement(20822..20916)  
/rpt\_family="MIR1J2"



from the RPCT-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBAC3.6  
This sequence is the entire insert of clone RP23-58B7.

FEATURES  
Location/Qualifiers

source

1. 241432  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="15"  
/clone="RP23-58B7"  
/clone\_lib="RPCT-23"  
147938..148190  
/note="Sequence from uni-directional dGTP big dye  
terminator reads only."

misc-feature  
BASE COUNT 64293 a 50995 c 55888 g 70256 t  
ORIGIN

Query Match 2.7%; Score 53.8; DB 10; Length 241432;  
Best Local Similarity 47.4%; Pred. No. 0.0024;  
Matches 228; Conservative 0; Mismatches 247; Indels 6; Gaps 2;

QY 305 AGGAGGAGGAATGCTTGGATCGAGATTCTTAAGAATGACTGGACATGTCAGAG 364  
|||||  
Db 147752 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147811  
QY 365 CCAGCTTCCCGAAGAACAGGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 424  
|||||  
Db 147812 AACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147871  
QY 425 GGGATACGCTGGAAGAACGCAATGCTACTGTGTATCTTGACAGCAGCCTTGCGCAGG 484  
|||||  
Db 147872 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147931  
QY 485 CCGAGATGCTGTGCTCCACACTGAAAAAGCATGAAGTACTTAGACGACGACGAGATG 544  
|||||  
Db 147932 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147991  
QY 545 AGACCAACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 604  
|||||  
Db 147992 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148045  
QY 605 AGATTGAGCTTCTACTCAGAGCCAGCTCCCTGAGTGGAGAGATGATCCGAGACATGG 664  
|||||  
Db 148046 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148105  
QY 665 GTGTGGAGCAGTCAGCGGTGGAACAGCTGCTGTGTACTGTGTCTCTCAAGAAAGAT 724  
|||||  
Db 148106 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148165  
QY 725 ACGGATCTTAAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 784  
|||||  
Db 148166 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148225  
QY 785 A 785  
Db 148226 A 148226

RESULT 39  
AC051623/c 195294 bp DNA linear HTG 06-SEP-2000  
LOCUS Mus musculus clone RP23-41205, WORKING DRAFT SEQUENCE, 17 unordered  
DEFINITION pieces.  
ACCESSION AC051623  
VERSION AC051623.4 GI:9972322  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 195294)

AUTHORS

McCombie,W.R., Baker,J.P., Bahret,A., Bal,H., Dedhia,N.N., de la  
Bastide,M., Huang,E.N., King,L., Kirchoff,K.A., Miller,B.,  
Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R., Rodriguez,M.A.,  
Shah,R.S., Shekher,M., Spiegel,L.A., Tolt,K. and Vill,M.D.  
Mouse genomic Sequence  
Unpublished

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 195294)  
McCombie,W.R.  
Direct Submission  
Submitted (15-APR-2000) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
On Sep 6, 2000 this sequence version replaced gi:8493559.

COMMENT

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor  
Laboratory  
Center code: CSHL  
Web site: <http://www.cshl.org/genseq>  
Contact: mcombie@cshl.org  
----- Project Information  
Center project name: RP23-41205  
Center clone name: RP23-41205

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 54595: contig of 54595 bp in length  
54596 54735: gap of unknown length  
54736 79339: contig of 24604 bp in length  
79340 79479: gap of unknown length  
79480 97436: contig of 17957 bp in length  
97437 97576: gap of unknown length  
97577 114331: contig of 16755 bp in length  
114332 114471: gap of unknown length  
114472 127452: contig of 12981 bp in length  
127453 127591: gap of unknown length  
127592 138079: contig of 10488 bp in length  
138080 138218: gap of unknown length  
138219 148562: contig of 10344 bp in length  
148563 148701: gap of unknown length  
148702 157987: contig of 9286 bp in length  
157988 158126: gap of unknown length  
158127 167394: contig of 9268 bp in length  
167395 167533: gap of unknown length  
167534 174603: contig of 7070 bp in length  
174604 174742: gap of unknown length  
174743 178566: contig of 3824 bp in length  
178567 178705: gap of unknown length  
178706 182160: contig of 3455 bp in length  
182161 182299: gap of unknown length  
182300 185278: contig of 2979 bp in length  
185279 185417: gap of unknown length  
185418 188252: contig of 2835 bp in length  
188253 188391: gap of unknown length  
188392 191030: contig of 2633 bp in length  
191031 191169: gap of unknown length  
191170 193703: contig of 2534 bp in length  
193704 193842: gap of unknown length  
193843 195294: contig of 1452 bp in length.

FEATURES  
Location/Qualifiers

source

1. 195294  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-41205"

BASE COUNT 48944 a 48556 c 46839 g 48701 t 2254 others  
ORIGIN

Query Match	2.7%	Score 53.4	DB 2	Length 195244
Best Local Similarity	49.5%	Pred. No. 0.003		
Matches 138	Conservative 0	Mismatches 141	Indels 0	Gaps 0
507	GAAGAAAGCATGATGACTTACTAGACGACGACGATGTGACCAACAAACAGCAAGAGGA	566		
55445	GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	55386		
567	GGCGGGCCCGCTTCAGGACGACAGATGAGCCATGTGACGACGATTGACCTTCTACTCCAGAG	626		
55385	GAGGAAAGAGGAAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	55326		
627	CCAGTCCCTGAGTGGGAGGAGATGATGCCAGACATGGGTGGGACATCGACGGGTGGA	686		
55325	GAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	55266		
687	ACAGCTGGCTGTGTCTGTCTGTCTGTCTCTCAAGAAAGATGTACGAAATCTTAAAGAGGACAG	746		
55265	GGAGCAG	55206		
747	GAGGCTTCAGGGGAGGTGGCTTGACACGCTGAGGAAGA	785		
55205	GCAG	55167		

RESULT	40
AC016961	202010 bp DNA linear HTG 13-DEC-2001
LOCUS	Human sapiens chromosome 3 clone RP11-394J21, WORKING DRAFT
DEFINITION	SEQUENCE, 6 unordered pieces.
ACCESSION	AC016961
VERSION	AC016961.16 GI:17530684
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 202010)
AUTHORS	Munzy,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C., Albrook,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Blamege,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,L.M., Davis,C., Dean-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denny,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Eliahi,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,C., Liu,W., Loulseged,H., Lozaco,R.J., Lu,X., Lucier,A., Luciger,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzner,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nokenkwo,S., Oguh,M., Okwoudu,G., Otagunye,N., Oyedero,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Prikmus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokun,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostlati,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tameitisa,A., Tameitisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.

```

TITLE
JOURNAL REFERENCE AUTHORS TITLE JOURNAL
Unpublished
2 (bases 1 to 202010)
Direct Submission Submitted (09-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 12, 2001 this sequence version replaced gi:117438263A.

COMMENT
----- genome center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMVB
Center clone name: RP11-394J21
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-primer Bodipy: 2% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 204811 bases at least Q40
Consensus quality: 206552 bases at least Q30
Consensus quality: 207618 bases at least Q20
Estimated insert size: 203982; sum-of-contigs estimation
Quality coverage: 8.1x In Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
be preserved.
    *   1      143460: contig of 143460 bp in length
    *       143461     gap of unknown length
    *       143561     gap of 21038 bp in length
    *       164598     gap of unknown length
    *       164699     gap of 13806 bp in length
    *       178504     contig of 13806 bp in length
    *       178505     gap of unknown length
    *       178605     contig of 10040 bp in length
    *       188644     gap of unknown length
    *       188645     contig of 10217 bp in length
    *       188745     gap of unknown length
    *       198961     gap of unknown length
    *       199062     contig of 2949 bp in length.
Location/Qualifiers
    1..202010
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="3"
        /clone="RP11-394J21"
BASE COUNT          59923 a 43696 c 42411 g 55465 t         515 others
ORIGIN
Query Match           2.7%; Score 53.4; DB 2; Length 202010;
Matches Local Similarity 49.5%; Pred.No.0.003;
Mismatches 138; Conservative 0; Mismatched 141; Indels 0; Gaps 0;

OY      507 GAATAAGCAGTGTGACTTAGACAGCACAGATGACCACAACAGCAAGAAGA 566
            |  |||  ||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 201656 GGAGAAAGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 201715
OY      567 GCCGGCGCCGCCTCAGAGCAGATTGAAACCATTCAGAGCAGATTCTACTCCAGAG 626
            |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```



**THIS PAGE BLANK (USPTO)**